(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 April 2003 (10.04.2003)

PCT

(10) International Publication Number
WO 03/029451 A2

(51) International Patent Classification[7]: C12N 9/16, 15/55, 5/10, C12Q 1/34, 1/68, C07K 16/40, G01N 33/573, A61K 38/46, 48/00, A61P 35/00

(21) International Application Number: PCT/GB02/04455

(22) International Filing Date: 2 October 2002 (02.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0123664.5 2 October 2001 (02.10.2001) GB

(71) Applicant (*for all designated States except US*): **THE INSTITUTE OF CANCER RESEARCH** [GB/GB]; 123 Old Brompton Road, London, Greater London SW7 3RP (GB).

(71) Applicants and
(72) Inventors: **ZELENT, Arthur** [US/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB). **PETRIE, Kevin** [GB/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB). **GUIDEZ, Fabien** [FR/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB).

(74) Agents: **KIDDLE, Simon, J.** et al.; Mewburn Ellis, York House, 23 Kingsway, London, Greater London WC2B 6HP (GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HISTONE DEACETYLASE 9

(57) Abstract: The identification and cloning of histone deacetylase 9 (HDAC9) is disclosed, and in particular full length HDAC9 polypeptides and HDAC9 polypeptides which have deacetylase activity, and to nucleic acid molecules encoding these polypeptides. The uses of these polypeptides and nucleic acid molecules are disclosed, for example for screening for compounds that are capable of modulating a HDAC9 biological activity.

WO 03/029451 A2

Histone Deacetylase 9

Field of the Invention

The present invention relates to the identification and isolation of a histone deacetylase and uses of this enzyme. In particular, the present invention relates to histone deacetylase 9 (HDAC9), a member of the class II histone deacetylase family.

Background of the Invention

Eukaryotic DNA is packaged into chromatin in a precise and highly regulated manner. This level of organisation is fundamental to many processes in the cell including replication, repair recombination, chromosomal segregation and transcriptional regulation. DNA is wrapped around the core histone octamer (H2A, H2B, H3 and H4) to form nucleosomes, which are the basic repeating units of chromatin. The crystal structure of the nucleosome has been solved and this has provided much information but there is still a great deal to learn about the mechanisms by which distinct functional domains of chromatin are formed and maintained. The most important changes in chromatin structure are thought to be influenced by post-translational modifications of N-terminal tails of the histones, which protrude from the nucleosome. These are highly basic when unmodified and interact with negatively charged DNA phosphate backbone. Specific amino acids within these tails are targets for a variety enzymes which can produce diverse modifications such as acetylation, methylation, phosphorylation, poly(ADP-ribosylation), and ubiquination. Acetylation is thus far the most widely studied and is catalysed by histone acetyltransferases (HATs) and involves substitution of the ε-amino group of specific lysines.

This leads to a more acidic residue and an overall decreased affinity for DNA by the histone octamer. It appears that the histone tails also mediate interactions between nucleosomes to form higher order chromatin structures and they could be disrupted by acetylation. The packaging of DNA into nucleosomal arrays presents a major physical obstacle to the transcriptional machinery when trying to gain access to the DNA template and there is strong evidence that unwinding of nucleosomes due to the acetylation of histone tails plays a fundamental role in the activation of gene transcription. In contrast to histone acetylation in transcriptional activation, enzymes which remove these modifications would be expected to have an important role in down-regulation and gene silencing. This has indeed been shown to be the case and recent studies have also implicated abnormal histone deacetylase function in a number of human cancers.

To date eight histone deacetylases have been characterised which may broadly divided into two related classes which share homology through their deacetylase domains with yeast histone deacetylases Rpd3 and Hda1. Class II HDACs 4, 5 and 7 may be differentiated from the Class I HDACs as they contain an additional N-terminal, non catalytic, region which is homologous to a protein previously characterized as a co-repressor, MEF2-Interacting Transcription Repressor (MITR/HDRP).

Zhou et al (PNAS, 98(19):10572-10577, 2001) purports to disclose the sequence of histone deacetylase HDAC9. However, the protein disclosed in the paper is incomplete and the level of deacetylase activity reported in the paper is below the negative control.

Summary of the Invention

Broadly, the present invention relates to the identification and cloning of histone deacetylase 9 (HDAC9), and in particular full length HDAC9 polypeptides and HDAC9 polypeptides which have deacetylase activity, and to nucleic acid molecules encoding these polypeptides. The present invention further relates to uses of these polypeptides and nucleic acid molecules, in particular for screening for compounds that are capable of modulating a HDAC9 biological activity.

Accordingly, in a first aspect, the present invention provides an isolated HDAC9 polypeptide having the amino acid sequence set out in Figure 2. This figure shows the cDNA and predicted amino acid sequence of wild-type histone deacetylase 9 (HDAC9), a 1069 amino acid polypeptide. The HDAC9 sequence is overlaid with the sequence of MITR that shares a common N-terminal amino acid sequence with HDAC9 polypeptide up to amino acid 577 at the end of exon 11, but has a different exon 12. The MITR sequence is a 593 amino acids long.

The HDAC9 sequence reported in Zhou et al (2001) is identical to the sequence shown in Figure 2 up to amino acid Ser1008, but does not include the C-terminal 61 amino acids which are encoded by exons 24 and 25 as shown in Figure 2. Significantly, the polypeptide reported in Zhou et al substantially lacks histone deacetylase activity. In contrast, preferred polypeptides of the present invention possess a HDAC9 biological activity, and especially histone deacetylase activity, and include all or part of the amino acid sequence set out between amino acids 1009 and 1069 of Figure 2. Compared to other histone deacetylases, HDAC9 has 55% sequence identity to

HDAC4, 56% sequence identity to HDAC5 and 38% sequence identity to HDAC7.

In a further aspect, the present invention provides an isolated polypeptide having at least 80% amino acid sequence identity with the amino acid sequence set out in Figure 2. Preferred embodiments of the invention provide isolated polypeptides having at least 95% sequence identity with the sequence set out in Figure 2.

In a further aspect, the present invention provides an isolated polypeptide which is encoded by a nucleic acid sequence which is capable of hybridising under stringent conditions to the nucleic acid sequence set out in Figure 2, or a complementary sequence thereof, and in particular to the nucleic acid sequence comprising a sequence which is capable of hybridising to a nucleic acid encoding amino acids 1009 to 1069 of the HDAC9 polypeptide as set out in Figure 2, or a complementary sequence thereof.

In a further aspect, the present invention provides a substance which is a fragment, active portion or sequence variant of one of the above polypeptides. Preferred fragments and active portions comprise all or part of the HDAC9 sequence set out between amino acids 1009 and 1069 inclusive as shown in Figure 2. Sequence variants are defined below and include HDAC9 polypeptides in which one or more of the exons shown in Figure 2 are deleted, such as the variants in which exon 7, exon 12 and/or exon 15 are deleted.

In a further aspect, the present invention provides a HDAC9 polypeptide as defined above joined to a coupling partner.

In a further aspect, the present invention provides an isolated nucleic acid molecule encoding one of the above polypeptides, and complementary nucleic acid sequences thereof. The cDNA sequence of full length HDAC9 is shown in Figure 2. The present invention also includes nucleic acid molecules having greater than a 90% sequence identity with one of the above nucleic acid sequence. In other embodiments, the present invention relates to nucleic acid sequences which hybridise to the coding sequence set out in Figure 2, or a complementary sequence thereof, and more especially to nucleic acid sequences which hybridise to a nucleic acid sequence encoding amino acids 1009 to 1069 as set out in Figure 2, or a complementary sequence thereof.

In further aspects, the present invention provides an expression vector comprising one of the above nucleic acid operably linked to control sequences to direct its expression, and host cells transformed with the vectors. The present invention also includes a method of producing a histone deacetylase 9 polypeptide, or a fragment or active portion thereof, comprising culturing the host cells and isolating the polypeptide thus produced.

In a further aspect, the present invention provides a composition comprising one or more of the above polypeptides or nucleic acid molecules as defined herein.

In a further aspect, the present invention provides the use of a histone deacetylase 9 polypeptide as defined herein (including fragments, active portions or sequence variants), or a corresponding nucleic acid molecule, for screening for candidate compounds which (a) share a histone deacetylase 9 biological activity or (b) bind to the histone deacetylase 9 polypeptide or (c) inhibit a biological activity of a histone deacetylase 9 polypeptide. By way of example, screening can be carried out to find peptidyl or non-peptidyl mimetics or inhibitors of the HDAC9 polypeptides to develop as lead compounds in pharmaceutical research.

Thus, in one embodiment, the present invention provides a method of identifying a compound which is capable of modulating an activity of a histone deacetylase 9 polypeptide, the method comprising:

(a) contacting at least one candidate compound with a histone deacetylase 9 (HDAC9) polypeptide as defined herein under conditions in which the candidate compound and HDAC9 polypeptide are capable of interacting;

(b) determining in an assay for a HDAC9 activity whether the candidate compound modulates the activity; and,

(c) selecting a candidate compound which modulates an activity of the HDAC9 polypeptide.

In a further aspect, the present invention provides a method of identifying a compound which is capable of inhibiting histone deacetylase 9 (HDAC9) polypeptide, the method comprising:

(a) contacting at least one candidate compound and a HDAC9 polypeptide as defined herein in the presence of a substrate for HDAC9 under conditions in which the candidate compound, HDAC9 polypeptide and HDAC9 substrate are capable of interacting;

(b) determining whether the candidate compound inhibits the activity of the HDAC9 polypeptide in reacting with the substrate; and,

(c) selecting the candidate compound which inhibits the activity of the HDAC9 polypeptide on the substrate.

In one embodiment, the substrate is a histone which is a substrate for HDAC9. In this assay, the effect of a candidate compound in modulating the activity of HDAC9 can be assessed by determining the amounts of acetylated and deacetylated histone present after the reaction of HDAC9 in the presence or absence of the candidate compounds. This can readily be determined using techniques well known in the art, for example employing antibodies which are capable of specifically binding to acetylated or deacetylated substrate.

In this aspect of the invention, preferably the activity of the HDAC9 polypeptide is the activity of removing an acetyl groups from the substrate. Conveniently, the progress of this reaction can be assessed by labelling the substrate with a detectable label (e.g. a radioactive label) and measuring the amount of label released from the substrate by the action of the HDAC9 polypeptide, e.g. in a scintillation proximity assay. Preferably, the method is for screening for modulators of HDAC9 which may be useful in for the treatment of cancer.

In a further aspect, the present invention provides a method of identifying a compound which is capable of inhibiting histone deacetylase 9 (HDAC9) polypeptide, the method comprising:

(a) contacting at least one candidate compound and a HDAC9 polypeptide as defined herein in the presence of a substrate for HDAC9 under conditions in which the candidate compound, HDAC9 polypeptide and HDAC9 substrate are capable of interacting;

(b) determining whether the candidate compound inhibits the activity of the HDAC9 polypeptide in reacting with the substrate; and,

(c) selecting the candidate compound which inhibits the activity of the HDAC9 polypeptide on the substrate.

In one embodiment, candidate compounds are screened for activity in modulating HDAC9 in a cell based reporter assay using cells which produce HDAC9 fused to a nucleic acid binding domain such as GAL4, e.g. by transfecting cells with appropriate expression and reporter vectors. In the cells, HDAC9-GAL4 chimeric protein which is capable of interacting with a GAL4 DNA binding site can bind to a promoter which includes such a site, the promoter being linked to a reporter construct. The HDAC9 activity can thus deacetylate core histones associated with a DNA surrounding GAL4 binding site resulting in down regulation of promoter activity. Compounds that inhibit HDAC9 activity would counteract this down regulation. Therefore, changes in reporter activity can be used as a screen for compounds that either stimulate or inhibit HDAC9 activity.

In a further aspect, the present invention provides antibodies capable of specifically binding to the above HDAC9 polypeptides, or an active portion, domain or fragment thereof, and the use of the HDAC9 polypeptide or peptides based on the sequence for designing antibodies or for use in a method of preparing antibodies capable of binding to HDAC9. These antibodies can be used in assays to detect and quantify the presence of HDAC9 polypeptide, in methods of purifying HDAC9 polypeptides, and as inhibitors of HDAC9 polypeptides.

In a further aspect, the present invention provides a method of amplifying a nucleic acid test sample comprising priming a nucleic acid polymerase reaction with nucleic acid encoding a HDAC9 polypeptide as defined above.

In a further aspect, the present invention provides a method for the diagnosis or prognosis of cancer, the method comprising determining the presence or amount of HDAC9 polypeptide, or an isoform thereof, or HDAC9 nucleic acid in a sample from a patient. This is discussed in more detail below.

These and other aspects of the present invention are described in more detail below. By way of example and not limitation, embodiments of the present invention will now be described in more detail with reference to the accompanying figures.

Brief Description of the Figures

Figure 1. Schematic representation of the rational behind the cloning of the 3' end HDAC9. EST accession nos. are indicated on the left with position numbers on the right. HDAC9 sequence is shown in bold and highlighted in grey. Sequence identity is indicated by vertical lines. The EST database at the NCBI was screened with EST sequences corresponding to HDAC9 cDNA found through homology to HDAC5. A number of overlapping clones were obtained which enabled the cloning of HDAC9 through the design of specific primers.

Figure 2. HDAC9 cDNA sequence together with the predicted HDAC9 amino acid sequence. HDAC9 full length cDNA has an open reading frame of 3207bp which yields a 1069 amino

acids protein. The HDAC9Δexon7 isoform results in an A to E substitution at position 222. The sequence specific to HDAC9ΔCD exon 12 is shown in italics.

Figure 3. Schematic representation of HDAC9 isoforms. The length of the various HDAC9 isoforms together with the primers used in cloning and RT-PCR analysis.

Figure 4. Analysis of HDAC9 sequence.

A. Amino acid sequence alignment of HDAC9, HDAC4, HDAC5, HDAC7, and a bacterial deacetylase, HDLP. The indicated sequences were aligned using Clustal W. Identical residues are boxed and highlighted in dark grey; similar residues are shaded in light grey.

B. Evaluation of amino acid identities and similarities of HDAC9 compared with other histone deacetylases. Values were obtained by comparing the whole protein or deacetylase domain of HDAC9 with the indicated protein sequences on BioEdit Sequence Alignment Editor using the Blosum62 matrix.

C. Phylogenetic tree of HDAC1 through to HDAC9. Sequences were aligned using the Clustal W server at the Centre for Molecular and Biomolecular Informatics (University of Nijmegen). The PHYLIP (bracket) notation output was used to construct an unrooted tree.

Figure 5. HDAC9 possesses HDAC activity and represses basal transcription.

A. HDAC9 deacetylates histone H4 peptide. 293T cells were transfected with FLAG-tagged HDAC. Whole cell lysates were produced and the HDACs precipitated with anti-FLAG agarose. The beads were assayed for their ability to deacetylate a [^{3}H]acetyl-labelled peptide corresponding to the N-Terminus of histone H4. Free acetate was extracted and measured by scintillation counting.

B. GAL4-HDAC9 Inhibits promoter activity *in-vivo* in a TSA sensitive manner. GAL4-Tk-Luc was transiently transfected into 293T cells together with 100ng GAL4DBD fusions. All Results were performed in duplicate and error.

Figure 6. Northern blot showing the differential expression of HDAC9 polypeptide in human tissues. A multiple human tissue Northern Blot was analysed with a ^{32}P labelled probe corresponding to the sequence found only in the full length HDAC9 cDNA in order to evaluate levels of expression of HDAC9 mRNA. The blot was reprobed with β-actin cDNA as a normalisation control. The tissues examined are indicated at the top of each lane. Positions of the RNA size markers are indicated in kilobases on the left of the blot.

Figure 7. HDAC9 displays higher activity towards histone H3 peptide. 293T cells were transfected with FLAG-tagged HDAC. Whole cell lysates were produced and the HDACs precipitated with anti-FLAG agarose. The beads were assayed for their ability to deacetylate a [^{3}H]acetyl-labelled peptide corresponding to the N-Terminus of Histone H3. Free acetate was extracted and measured by scintillation counting.

Figure 8. Expression of various HDAC9 isoforms in normal and cancer cells. Total RNA was isolated from the indicated cell lines and tissues and cDNA generated with

M-MLV reverse transcriptase using either gene specific or random primers. Standard PCR was performed for 32 cycles.

Figure 9. In vitro interactions of HDAC9 with various oncoproteins and co-repressors GST-HDAC9, GST-HDAC9CD and GST-MITR were produced from *E. coli* DH5α. [^{35}S]methionine-labeled proteins were synthesized *in-vitro* using a rabbit reticulocyte lysate-coupled transcription-translation system. Assays were performed in NETN Buffer, washed in H Buffer, separated by SDS-PAGE and visualised by autoradiography.

Figure 10. Cellular localization of various isoforms of HDAC9. COS-1 cells were transiently transfected with N-Terminus flag tagged HDAC9, HDAC9ΔCD/MITR, HDAC9ΔExon7, HDAC9 ΔExon12 or HDAC9ΔExon15 as indicated. After methanol fixation, cells were stained with dapi (blue) and antibodies raised against HDAC9 or flag (green). G-15 is an antibody to the N-Terminus of HDAC9 which h detects all isoforms (including HDAC9ΔCD/MITR).

Figure 11. Effect of various isoforms of HDAC9 on the cellular localization of BCL-6. COS-1 cells were transiently transfected with N-Terminus flag tagged HDAC9, HDAC9ΔCD, or HDAC9Δexon7, as indicated. After methanol fixation, cells were stained with dapi (blue) and antibodies raised against HDAC9 G-15 (green) or BCL-6(red).

Figure 12. Effect of various isoforms of HDAC9 on the cellular localization of NCoR. COS-1 cells were transiently transfected with N-Terminus flag tagged HDAC9, HDAC9ΔCD, or HDAC9Δexon7, as indicated After

methanol fixation, cells were stained with dapi (blue) and antibodies raised against FLAG (green) G-15 or NCoR (red).

Figure 13. Effect of various isoforms of HDAC9 on the cellular localization of PLZF. COS-1 cells were transiently transfected with N-Terminus flag tagged HDAC9, or HDAC9 Δexon7, as indicated After methanol fixation, cells were stained with dapi (blue) and antibodies raised against HDAC9 G-15 (green) or PLZF (red).

Histone Deacetylase 9 Nucleic Acid

A "HDAC9 nucleic acid" includes a nucleic acid molecule which has a nucleotide sequence encoding a polypeptide comprising the amino acid sequence shown in Figure 2 or any one of the other HDAC9 polypeptides of the present invention. The HDAC9 coding sequence may be the full length nucleic acid sequence shown in Figure 2, a complementary nucleic acid sequence, or it may be a sequence variant differing from one of the above sequences by one or more of addition, insertion, deletion and substitution of one or more nucleotides of the sequence shown. Changes to a nucleotide sequence may result in an amino acid change at the protein level, or not, as determined by the genetic code. Nucleic acid encoding a polypeptide which is a sequence variant preferably has at least 80% sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity, and most preferably at least 99% sequence identity with the nucleic acid sequence shown in Figure 2.

The present invention also includes fragments of the

HDAC9 nucleic acid sequences described herein, the fragments preferably being at least 60, 120, 180, 240, 480 or 960 nucleotides in length. Preferred HDAC9 nucleic acid fragments share at least part of their nucleic acid sequence with the HDAC9 nucleic acid sequence which encodes amino acids 1008 to 1069 of the polypeptide shown in Figure 2, or a complementary sequence thereof.

Generally, nucleic acid according to the present invention is provided as an isolate, in isolated and/or purified form, or free or substantially free of material with which it is naturally associated. Nucleic acid may be wholly or partially synthetic and may include genomic DNA, cDNA or RNA. Where nucleic acid according to the invention includes RNA, reference to the sequence shown should be construed as reference to the RNA equivalent, with U substituted for T.

The present invention also includes nucleic acid molecules which are capable of hybridising to one of the HDAC9 sequences disclosed herein, or a complementary sequence thereof. Stringency of hybridisation reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridisation generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridisable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridisation reactions, see Ausubel et al, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

Preferably, a nucleic acid sequence will hybridise to a HDAC9 sequence of the invention, or a complementary sequence thereof under "stringent conditions". These are well known to those skilled in the art and include those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/ 0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridisation a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 760 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6 8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

Nucleic acid sequences encoding all or part of the HDAC9 gene and/or its regulatory elements can be readily prepared by the skilled person using the information and references contained herein and techniques known in the art (for example, see Sambrook, Fritsch and Maniatis, Molecular Cloning, A Laboratory Manual, Cold Spring Harbour Laboratory Press, 1989, and Ausubel et al, Short Protocols in Molecular Biology, John Wiley and Sons, 1992). These techniques include (i) the use of the polymerase chain reaction (PCR) to amplify samples of such nucleic acid, e.g. from genomic sources, (ii) chemical synthesis, or (iii) amplification in *E. coli*. Modifications to the HDAC9 sequences can be made, e.g. using site directed mutagenesis, to provide expression of modified HDAC9 polypeptide or to take account of codon preference in the host cells used to express the nucleic acid.

PCR techniques for the amplification of nucleic acid are described in US Patent No. 4,683,195. The HDAC9 nucleic acid sequences provided herein readily allow the skilled person to design PCR primers. References for the general use of PCR techniques include Mullis et al, Cold Spring Harbour Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR Technology, Stockton Press, NY, 1989, Ehrlich et al, Science, 252:1643-1650, (1991), "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, Academic Press, New York, (1990).

In order to obtain expression of the HDAC9 nucleic acid sequences, the sequences can be incorporated in a vector having control sequences operably linked to the HADC9 nucleic acid to control its expression. The vectors may include other sequences such as promoters or enhancers to drive the expression of the inserted nucleic acid, nucleic acid sequences so that the HDAC9 polypeptide is produced as a fusion and/or nucleic acid encoding secretion signals so that the polypeptide produced in the host cell is secreted from the cell. Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. Vectors may be plasmids or viral, e.g. 'phage, or phagemid, as appropriate. For further details see, for example, Molecular Cloning: a Laboratory Manual: 2nd edition, Sambrook et al., 1989, Cold Spring Harbour Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in Current Protocols in Molecular Biology, Ausubel et al. eds., John Wiley & Sons, 1992.

HDAC9 polypeptide can be obtained by transforming the vectors into host cells in which the vector is functional, culturing the host cells so that the HDAC9 polypeptide is produced and recovering the HDAC9 polypeptide from the host cells or the surrounding medium. Prokaryotic and eukaryotic cells are used for this purpose in the art, including strains of *E. coli*, insect cells (e.g. transformed with baculovirus), yeast, and eukaryotic cells such as COS or CHO cells. The choice of host cell can be used to control the properties of the HDAC9 polypeptide expressed in those cells, e.g. controlling where the polypeptide is deposited in the host cells or affecting properties such as its glycosylation and phosphorylation. If the polypeptide is expressed coupled to an appropriate signal leader peptide it may be secreted from the cell into the culture medium. Following production by expression, a polypeptide may be isolated and/or purified from the host cell and/or culture medium, as the case may be, and subsequently used as desired, e.g. in the formulation of a composition which may include one or more additional components such as a carrier. Polypeptides may also be expressed in *in vitro* systems, such as reticulocyte lysate.

The nucleic acid sequences provided herein are useful for identifying HDAC9 nucleic acid in a test sample. The present invention provides a method of obtaining nucleic acid of interest, the method including hybridising a probe sharing all or part of the sequence provided herein, or a complementary sequence, to the target nucleic acid. Hybridisation is generally followed by identification of successful hybridisation and isolation of nucleic acid which has hybridised to the probe, which may involve one or more steps of PCR. These methods may be useful in determining whether HDAC9 nucleic acid is present in a sample, e.g. in a particular type of cells present in the sample.

Nucleic acid according to the present invention is obtainable using one or more oligonucleotide probes or primers designed to hybridise with one or more fragments of the nucleic acid sequence shown herein, particularly fragments of relatively rare sequence, based on codon usage or statistical analysis. A primer designed to hybridise with a fragment of the nucleic acid sequence shown in the above figures may be used in conjunction with one or more oligonucleotides designed to hybridise to a sequence in a cloning vector within which target nucleic acid has been cloned, or in so-called "RACE" (rapid amplification of cDNA ends) in which cDNA's in a library are ligated to an oligonucleotide linker and PCR is performed using a primer which hybridises with the sequence shown herein and a primer which hybridises to the oligonucleotide linker.

On the basis of amino acid sequence information, oligonucleotide probes or primers may be designed, taking into account the degeneracy of the genetic code, and where appropriate, codon usage of the organism from the candidate nucleic acid is derived. An oligonucleotide for use in nucleic acid amplification may have about 10 or fewer codons (e.g. 6, 7 or 8), i.e. be about 30 or fewer nucleotides in length (e.g. 18, 21 or 24). Generally specific primers are upwards of 14 nucleotides in length, but not more than 18-20. Those skilled in the art are well versed in the design of primers for use processes such as PCR.

Accordingly, a further aspect of the present invention provides an oligonucleotide or nucleotide fragment of the one of the nucleotide sequence disclosed herein, or a complementary sequence, in particular for use in a method of obtaining and/or screening nucleic acid. The sequences referred to above may be modified by addition, substitution, insertion or deletion of one or more nucleotides, but preferably without abolition of ability to hybridise selectively with nucleic acid with the sequence shown herein, that is wherein the degree of sequence identity of the oligonucleotide or polynucleotide with one of the sequences given is sufficiently high.

In some preferred embodiments, oligonucleotides according to the present invention that are fragments of any of the nucleic acid sequences provided herein, or complementary sequences thereof, are at least about 10 nucleotides in length, more preferably at least about 15 nucleotides in length, more preferably at least about 20 nucleotides in length. Such fragments themselves individually represent aspects of the present invention. Fragments and other oligonucleotides may be used as primers or probes as discussed but may also be generated (e.g. by PCR) in methods concerned with determining the presence of HDAC9 nucleic acid in a test sample.

HDAC9 Polypeptides

The HDAC9 polypeptides disclosed herein, or fragments or active portions thereof, can be used as pharmaceuticals, in the developments of drugs, for further study into its properties and role *in vivo* and to screen for HDAC9 inhibitors. Thus, a further aspect of the present invention provides a polypeptide which has the amino acid sequence shown in Figure 2, which may be in isolated and/or purified form, free or substantially free of material with which it is naturally associated, such as polypeptides other than HDAC9.

The present invention also includes active portions, domains and fragments (including domains) of the HDAC9 polypeptides of the invention.

An "active portion" of HDAC9 polypeptide means a peptide which is less than said full length HDAC9 polypeptide, but which retains at least some of its essential biological activity, e.g. as a deacetylase. Active portions may be great than 100 amino acids, more preferably greater than 200 amino acids, more preferably greater than 300 amino acids and most preferably greater than 400 amino acids in length. Preferred active portion include all or part of the 61 C-terminal amino acids shown in Figure 2.

A "fragment" of the HDAC9 polypeptide means a stretch of amino acid residues of at least 5 contiguous amino acids from the sequences set out as Figure 2, or more preferably at least 7 contiguous amino acids, or more preferably at least 10 contiguous amino acids or more preferably at least 20 contiguous amino acids or more preferably at least 40 contiguous amino acids. Fragments of the HDAC9 polypeptide sequences may be useful as antigenic determinants or epitopes for raising antibodies to a portion of the HDAC9 amino acid sequence which also forms part of the present invention. For instance, fragments of HDAC9 can act as sequestrators or competitive antagonists by interacting with other proteins, e.g. if they possess a protein interaction domain present in the full length HDAC9 sequence.

Polypeptides which are amino acid sequence variants are also provided by the present invention. A "sequence variant" of the HDAC9 polypeptide, or an active portion or fragment thereof, means a polypeptide modified by varying the amino acid sequence of the protein, e.g. by manipulation of the nucleic acid encoding the protein or by altering the protein itself. Such sequence variants of the natural amino acid sequence may involve insertion, addition, deletion or substitution of one, two, three, five, ten, twenty or more amino acids.

Preferred polypeptides have HDAC9 biological activity, that is, they are capable of removing an acetyl group from a substrate such as a histone (e.g. histone H3 or H4). HDAC9 and isoforms of the protein may have other activities and one potential role is linked with the

Δexon7 isoform. It has been suggested that HDACs can function as molecular reservoirs within the cell as well as deacetylating histones, and also as transcription factors. We have observed that HDAC9 interacts *in vitro* and co-localizes *in vivo* with a number of transcriptional repressors, such as TEL, PLZF and BCL-6, whose function has been implicated in the pathogenesis of hematological malignancies, as well as the co-repressors mSIN3A/B and NCoR. When HDAC9 or HDAC9Δexon7 are co-transfected into COS cells they have the ability to delocalize BCL-6 and PLZF. NCoR is also delocalised. The interesting aspect of these interactions is that HDAC9 is located in the nucleus whilst HDAC9Δexon7 is cytoplasmic. Additionally, the HDAC9ΔCD/hMITR isoform displays a microspeckled distribution within the nucleus which is distinct to that of HDAC9.

Preferably, a polypeptide which is an amino acid sequence variant of the amino acid sequence shown in Figure 2 at least 80% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least 97% sequence identity more preferably at least 98% sequence identity and most preferably at least 99% sequence identity to the sequences of Figure 2.

The skilled person can readily make sequence comparisons and determine identity using techniques well known in the art, e.g. using the GCG program which is available from Genetics Computer Group, Oxford Molecular Group, Madison, Wisconsin, USA, Version 9.1. Particular amino acid sequence variants may differ from those shown in SEQ ID Nos: 2 or 4 by insertion, addition, substitution or deletion of 1 amino acid, 2, 3, 4, 5-10, 10-20 20-30, 30-

50, 50-100, 100-150, or more than 150 amino acids.

"Percent (%) amino acid sequence identity" with respect to the HDAC9 polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the HDAC9 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. The % identity values can be generated by WU-BLAST-2 which was obtained from [Altschul et al, Methods in Enzymology, 266:460-480 (1996); http://blast.wustl/edu/blast/README.html]. WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span=1, overlap fraction= 0.125, word threshold(T)=11. The HSPS and HSPS2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity. A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the "longer" sequence in the aligned region. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored).

Similarly, "percent (%) nucleic acid sequence identity" with respect to the coding sequence of the HDAC9 polypeptides identified herein is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues in the HDAC9 coding sequence as provided in Figure 2. The identity values used herein were generated by the BLASTN module of WU BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively.

A polypeptide according to the present invention may be isolated and/or purified (e.g. using an antibody) for instance after production by expression from encoding nucleic acid. Polypeptides according to the present invention may also be generated wholly or partly by chemical synthesis. The isolated and/or purified polypeptide may be used in formulation of a composition, which may include at least one additional component.

The HDAC9 polypeptides can also be linked to a coupling partner, e.g. an effector molecule, a label, a drug, a toxin and/or a carrier or transport molecule. Techniques for coupling the peptides of the invention to both peptidyl and non-peptidyl coupling partners are well known in the art.

Antibodies Capable of Binding HDAC9 Polypeptides

A further important use of the HDAC9 polypeptides is in raising antibodies that have the property of specifically binding to the HDAC9 polypeptides or fragments thereof. The techniques for producing monoclonal antibodies to HDAC9 protein are well established in the art. Preferred antibodies are capable of specifically binding to an epitope of HDAC9 located or partially located between amino acids 1009 and 1069 of the amino acid sequence shown in Figure 2. The anti-HDAC9 antibodies of the present invention may be specific in the sense of being able to distinguish between the polypeptide it is able to bind and other human polypeptides for which it has no or substantially no binding affinity (e.g. a binding affinity of about 1000x worse). Specific antibodies bind an epitope on the molecule which is either not present or is not accessible on other molecules. Other preferred antibodies include those capable of substantially neutralising a HDAC9 biological activity and especially biological activity as a histone deacetylase.

Antibodies may be obtained using techniques which are standard in the art. Methods of producing antibodies include immunising a mammal (e.g. mouse, rat, rabbit, horse, goat, sheep or monkey) with the protein or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage et al, Nature, 357:80-82, 1992). Isolation of antibodies and/or antibody-producing cells from an animal may be accompanied by a step of sacrificing the animal. As an alternative or supplement to immunising a mammal with a peptide, an antibody specific for a protein may be obtained from a recombinantly produced library of expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047. The library may be naive, that is constructed from sequences obtained from an organism which has not been immunised with any of the proteins (or fragments), or may be one constructed

using sequences obtained from an organism which has been exposed to the antigen of interest.
Antibodies according to the present invention may be modified in a number of ways that are well known in the art. Indeed the term "antibody" should be construed as covering any binding substance having a binding domain with the required specificity. Thus the invention covers antibody fragments, derivatives, functional equivalents and homologues of antibodies, including synthetic molecules and molecules whose shape mimics that of an antibody enabling it to bind an antigen or epitope. Humanised antibodies in which CDRs from a non-human source are grafted onto human framework regions, typically with the alteration of some of the framework amino acid residues, to provide antibodies which are less immunogenic than the parent non-human antibodies, are also included within the present invention.

A hybridoma producing a monoclonal antibody according to the present invention may be subject to genetic mutation or other changes. It will further be understood by those skilled in the art that a monoclonal antibody can be subjected to the techniques of recombinant DNA technology to produce other antibodies or chimeric molecules which retain the specificity of the original antibody. Such techniques may involve introducing DNA encoding the immunoglobulin variable region, or the complementarity determining regions (CDRs), of an antibody to the constant regions, or constant regions plus framework regions, of a different immunoglobulin. See, for instance, EP 0 184 187 A, GB 2 188 638 A or EP 0 239 400 A. Cloning and expression of chimeric antibodies are described in EP 0 120 694 A and EP 0 125 023 A.
In a further aspect the present invention provides a

method of making antibodies, the method comprising employing a HDAC9 polypeptide or a fragment thereof as an immunogen. The present invention also provides a method of screening for antibodies which are capable of specifically binding HDAC9 polypeptide, the method comprising contacting a HDAC9 polypeptide with one or more candidate antibodies and detecting whether binding occurs.

Preferred antibodies according to the invention are isolated, in the sense of being free from contaminants such as antibodies able to bind other polypeptides and/or free of serum components. Monoclonal antibodies are preferred for some purposes, though polyclonal antibodies are within the scope of the present invention.

Hybridomas capable of producing antibody with desired binding characteristics are within the scope of the present invention, as are host cells, eukaryotic or prokaryotic, containing nucleic acid encoding antibodies (including antibody fragments) and capable of their expression. The invention also provides methods of production of the antibodies including growing a cell capable of producing the antibody under conditions in which the antibody is produced, and preferably secreted.

The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule. One favoured mode is by covalent linkage of each antibody with an individual fluorochrome, phosphor or laser exciting dye with spectrally isolated absorption or emission characteristics. Suitable fluorochromes include fluorescein, rhodamine, phycoerythrin and Texas Red. Suitable chromogenic dyes include diaminobenzidine.

Other reporters include macromolecular colloidal particles or particulate material such as latex beads that are coloured, magnetic or paramagnetic, and biologically or chemically active agents that can directly or indirectly cause detectable signals to be visually observed, electronically detected or otherwise recorded. These molecules may be enzymes which catalyse reactions that develop or change colours or cause changes in electrical properties, for example. They may be molecularly excitable, such that electronic transitions between energy states result in characteristic spectral absorptions or emissions. They may include chemical entities used in conjunction with biosensors. Biotin/avidin or biotin/streptavidin and alkaline phosphatase detection systems may be employed.

Antibodies according to the present invention may be used in screening for the presence of a polypeptide, for example in a test sample containing cells or cell lysate as discussed, and may be used in purifying and/or isolating a polypeptide according to the present invention, for instance following production of the polypeptide by expression from encoding nucleic acid. Antibodies may modulate the activity of the polypeptide to which they bind and so, if that polypeptide has a deleterious effect in an individual, may be useful in a therapeutic context (which may include prophylaxis).

Screening for Mimetic Substances

The present invention further relates to the use of a histone deacetylase 9 polypeptide or nucleic acid molecule for screening for candidate compounds which (a) share a histone deacetylase 9 biological activity or (b) bind to the histone deacetylase 9 polypeptide or (c) inhibit a biological activity of a histone deacetylase 9 polypeptide.

It is well known that pharmaceutical research leading to the identification of a new drug may involve the screening of very large numbers of candidate substances, both before and even after a lead compound has been found. This is one factor which makes pharmaceutical research very expensive and time-consuming. Means for assisting in the screening process can have considerable commercial importance and utility.

By way of example, screening can be carried out to find peptidyl or non-peptidyl mimetics or inhibitors of the HDAC9 polypeptides to develop as lead compounds in pharmaceutical research.

In this aspect of the invention, preferably the activity of the HDAC9 polypeptide is the activity of removing an acetyl groups from the substrate. Conveniently, the progress of this reaction can be assessed by labelling the substrate with a detectable label (e.g. a radioactive label) and measuring the amount of label released from the substrate by the action of the HDAC9 polypeptide, e.g. in a scintillation proximity assay. Preferably, the method is for screening for modulators of HDAC9 which may be further tested for use a therapeutic, especially for the treatment of cancer.

A method of screening for a substance which modulates activity of a polypeptide may include contacting one or more test substances with a HDAC9 in a suitable reaction medium, testing the activity of the treated polypeptide and comparing that activity with the activity of the polypeptide in comparable reaction medium untreated with the test substance or substances. A difference in activity between the treated and untreated polypeptides is indicative of a modulating effect of the relevant test substance or substances.

Combinatorial library technology provides an efficient way of testing a potentially vast number of different substances for ability to modulate activity of a polypeptide. Such libraries and their use are known in the art. The use of peptide libraries is preferred.

Prior to or as well as being screened for modulation of activity, test substances may be screened for ability to interact with the polypeptide, e.g. in a yeast two-hybrid system (which requires that both the polypeptide and the test substance can be expressed in yeast from encoding nucleic acid). This may be used as a coarse screen prior to testing a substance for actual ability to modulate activity of the polypeptide. Alternatively, the screen could be used to screen test substances for binding to a HDAC9 specific binding partner, to find mimetics of the HDAC9 polypeptide, e.g. for testing as therapeutics.

In one embodiment, the present invention provides a method of identifying a compound which is capable of modulating an activity of a histone deacetylase 9 polypeptide, the method comprising:

(a) contacting at least one candidate compound with a histone deacetylase 9 (HDAC9) polypeptide as defined herein under conditions in which the candidate compound and HDAC9 polypeptide are capable of interacting;

(b) determining in an assay for a HDAC9 activity whether the candidate compound modulates the activity; and

(c) selecting a candidate compound which modulates an activity of the HDAC9 polypeptide.

In a preferred embodiment, the present invention provides a method of identifying a compound which is capable of inhibiting histone deacetylase 9 (HDAC9) polypeptide, the method comprising:

(a) contacting at least one candidate compound and a HDAC9 polypeptide as defined herein in the presence of a substrate for HDAC 9 under conditions in which the candidate compound, HDAC9 polypeptide and HDAC9 substrate are capable of interacting;

(b) determining whether the candidate compound inhibits the activity of the HDAC9 polypeptide in reacting with the substrate; and,

(c) selecting the candidate compound which inhibits the activity of the HDAC9 polypeptide on the substrate.

Following identification of a candidate compound which modulates or inhibits HDAC9 activity, the substance may be investigated further. Furthermore, it may be manufactured and/or used in preparation, i.e. manufacture or formulation, of a composition such as a medicament, pharmaceutical composition or drug.

Diagnostic Methods

Recent studies have implicated abnormal histone deacetylase function with a number of human cancers. Accordingly, the present invention also provides the use of HDAC9 as a diagnostic marker for cancer, e.g. by correlating this level with the amount of the HDAC9 polypeptide, an isoform thereof, or HDAC9 nucleic acid present in a control.

Under normal circumstances, the expression of HDAC9 isoforms which contain the catalytic domain appears to be tightly regulated, with the protein only found at high levels in adult and foetal brain, and to a lesser extent skeletal muscle, testes, bone marrow, thymus, spleen, $CD14^{+ve}$, and $CD19^{+ve}$ cells. However, inspection of the RT PCR data reveals that HDAC9 is specifically expressed in TEL-AML1 positive and negative pre-B cell acute lymphoblastic leukaemia, and B cell lymphoma cell lines and patient samples. HDAC9 is not expressed in various acute myeloid leukaemia cell lines, with the exception of acute monocytic leukaemia, possibly reflecting a common progenitor for B cell and monocytic lineages. HDAC9 is also found at lower levels in the erythroleukemia cell line, HEL. The cell line which overexpress HDAC9 also express the ΔExon7 and ΔExon12 isoforms at higher levels and in one the of CLL patient samples which have been stimulated with SAC/IL2, there is a clear change in the relative expression of full length and ΔExon12 isoforms. There is evidence that the addition of HDAC inhibitors can alter the cellular distribution of HDACs presumably through interfering with protein-protein interactions as well blocking the catalytic site and this potential of HDAC9 in recruiting and modifying of other proteins, as well histone deacetylation. Taking this into consideration, the ΔExon7 may prove to be very important in the pathogenesis of some leukaemias.

In this context, there are a number of methods known in the art for analysing samples from individuals to determine the presence of HDAC9 nucleic acid or polypeptide. The assays may determine the presence or amount of HDAC9 nucleic acid or polypeptide in a sample from a patient, and whether the nucleic acid or polypeptide is full length or has a HDAC9 biological activity. Examples of biological samples include blood, plasma, serum, tissue samples, tumour samples, saliva and urine. The purpose of such analysis may be used for diagnosis or prognosis, to assist a physician in determining the severity or likely course of the condition and/or to optimise treatment of it.

Exemplary approaches for detecting HDAC9 nucleic acid or polypeptides include:

(a) determining the presence or amount of HDAC9 polypeptide in a sample from a patient, by measuring an activity of the HDAC9 polypeptide or its presence in a binding assay; or,

(b) determining the presence of HDAC9 nucleic acid using a probe capable of hybridising to the HDAC9 nucleic acid;

(c) using PCR involving one or more primers based on a HDAC9 nucleic acid sequence to determine whether the HDAC9 transcript is present in a sample from a patient.

In one embodiment, the method comprises the steps of:

(a) contacting a sample obtained from the patient with a solid support having immobilised thereon binding agent having binding sites specific for HDAC9 polypeptide or HDAC9 nucleic acid;

(b) contacting the solid support with a labelled developing agent capable of binding to unoccupied binding sites, bound HDAC9 polypeptide or nucleic acid or occupied binding sites; and,

(c) detecting the label of the developing agent specifically binding in step (b) to obtain a value representative of the presence or amount of the HDAC9 polypeptide or nucleic acid in the sample.

The binding agent preferably is a specific binding agent and has one or more binding sites capable of specifically binding to HDAC9 or nucleic acid in preference to other molecules. Conveniently, the binding agent is immobilised on solid support, e.g. at a defined location, to make it easy to manipulate during the assay.

Examples of specific binding pairs are antigens and antibodies, molecules and receptors and complementary nucleotide sequences. The skilled person will be able to think of many other examples and they do not need to be listed here. Further, the term "specific binding pair" is also applicable where either or both of the specific binding member and the binding partner comprise a part of a larger molecule. In embodiments in which the specific binding pair are nucleic acid sequences, they will be of a length to hybridise to each other under the conditions of the assay, preferably greater than 10 nucleotides long, more preferably greater than 15 or 20 nucleotides long.

There are various methods for determining the presence or absence in a test sample of a particular nucleic acid sequence, such as the sequence shown in Figure 2.

Exemplary tests include nucleotide sequencing, hybridisation using nucleic acid immobilized on chips, molecular phenotype tests, protein truncation tests (PTT), single-strand conformation polymorphism (SSCP) tests, mismatch cleavage detection and denaturing gradient gel electrophoresis (DGGE). These techniques and their advantages and disadvantages are reviewed in Nature Biotechnology, 15:422-426, 1997.

Pharmaceutical Compositions

The present invention disclose the use of HDAC9 or inhibitors thereof for formulation in pharmaceutical compositions, and especially compositions for the treatment of cancer, and more especially leukaemias such as TEL-AML1 positive and negative pre-B cell acute lymphoblastic leukaemia, and B cell lymphoma. These compositions may comprise, in addition to one of the above substances, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material may depend on the route of administration, e.g. oral, intravenous, cutaneous or subcutaneous, nasal, intramuscular, intraperitoneal routes.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may include a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally include a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Whether it is a polypeptide, antibody, peptide, nucleic acid molecule, small molecule or other pharmaceutically useful compound according to the present invention that is to be given to an individual, administration is preferably in a "prophylactically effective amount" or a "therapeutically effective amount" (as the case may be, although prophylaxis may be considered therapy), this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, e.g. decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors, and typically takes account of the disorder to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 20th Edition, 2000, pub.

Lippincott, Williams & Wilkins. A composition may be administered alone or in combination with other treatments, either simultaneously or sequentially, dependent upon the condition to be treated.

Experimental

Cloning of HDAC9

The National Institute for Biotechnology Information (NCBI) high throughput genomic sequence database (htgs) was searched with the amino acid sequences corresponding to the deacetylase domain of HDAC5 in an effort to identify novel histone deacetylase genes. Several DNA sequences encoding peptides with significant homology to HDAC5 were found on a human BAC clone RP11-8I15 (accession no. AC016186) which contains 70 unordered contigs and maps to chromosome 18. When the Genbank nucleotide database was searched with a composite of the novel sequences showing homology to HDAC5, it aligned with BAC clone CTB-13P7 (accession no. AC002088) which mapped to chromosome 7p15-p21 (Table 1), indicating the presence of a pseudogene or incorrect entry into the htgs database. A search of the literature revealed that 7p15-p21 had already been identified as containing a potential HDAC open reading frame and that a gene encoding a protein called HDRP/MITR which shares 50% identity with the N-terminus of HDACs 4 and 5, had been cloned and also mapped to this genomic region. Since a search of the expressed sequence tag database (dbEST) revealed that cDNA corresponding to a partial HDAC domain had been isolated from germinal centre B cells (accession no. AA287983), several hematopoietic cell lines were analysed by reverse transcription (RT)-PCR with oligonucleotides corresponding to the HDAC domain and HDRP/MITR. It was found that not only was it possible to amplify bands from the HDAC domain itself but also from the cDNA of HDRP/MITR through to the HDAC domain and this resulted in the generation of around 2700 bp of sequence. 3′ RACE failed to yield the remaining sequence so an attempt was made to try to find overlapping ESTs in order to 'walk' along the cDNA (Figure 1) and confirm any matching sequence against the genomic clones. During this investigation, it became apparent that the sequence corresponding to the final 5 exons of HDAC9 (accession no. RP5-1194E15 and GS1-465N13) had been submitted in the antisense direction in relation to the rest of the gene (see Table 1). As these clones accounted for almost 160kb of DNA (out of a total of 500kb for the entire gene), search algorithms had not previously identified homology between the EST and genomic sequences. Clones spanning the HDAC9 gene are listed in Table 2. When the database searches were repeated with the last two genomic clones in the correct orientation, an alignment was obtained, which showed significant homology with the sequence for HDACs 4 and 5. This information permitted the cloning of the entire open reading frame of HDAC9 (Figure 2) from Marathon-Ready Human Brain cDNA (Clontech) using the sense primer 9F1 5'-ATGCACAGTATGATCAGCTCA-3' and the antisense primer 9R1 5'-GTCACACACAGGAAATATCAG-3'. HDAC9ΔCD/HDRP/MITR was cloned from Marathon-Ready Human Brain cDNA (Clontech) using the sense primer 9F1 and the antisense primer 9ΔCDR1 5'-TCAGATAATGACTTTAATTACAAAT-3' (see Figure 2 for the cDNA sequence and translation). HDAC9ΔExon7 and HDAC9ΔExon12 were cloned from the acute monocytic leukaemia cell line MONO-MAC-6 using the sense primer 9F1 and the antisense primer 9R3 5'-TCTCTAATCCATCCATGCCAA-3'. HDAC9ΔExon15 was cloned from the acute Pre-B ALL cell line REH using the sense primer 9F2 5'-AGGCTGCTTTTATGCAACAG-3' and the antisense primer 9R2 5'-CTGAATGCTTCAAGGTACTCA-3'. See Figure 3 for a schematic detailing the positions of the above primers. PCR products were cloned into pCRII (Invitrogen) and sequenced using BigDye (Perkin Elmer).

Analysis of HDAC9 sequence

The full-length product of the HDAC9 gene comprises 1069 amino acids as shown in Figures 2 (see Figure 3 for schematic), and is encoded by exons 2-26 (exon 1 is untranslated). The 26 exons, which form the HDAC9 cDNA, span 500 kb of genomic sequence on chromosome 7p15-p21. HDAC9ΔCD, also referred to as HDRP/MITR is 593 amino acids in length and contains 16 residues of unique sequence encoded by a region of exon 12 which is 3' to the splice donor site used to generate the HDAC9 ORF. There are several exon deletions which may occur that preserve the open reading frame of HDAC9 (Table 2) and two have been identified and cloned. HDAC9Δexon7 is 1025 amino acids long and contains an Ala to Glu substitution at position 222 as the result of the deletion of exon 7. This isoform lacks two serines (S223 and S253) which when phosphorylated have been implicated in 14-3-3 protein dependent shuttling of HDAC4 and 5 from the nucleus to the cytoplasm, and also a tripartite nuclear localization signal. HDAC9Δexon12 is 981 amino acids long and lacks exon 12 sequences encoding a leucine zipper motif that may mediate interactions of HDAC9 with other proteins.

An isoform, which possesses neither exon 7 nor exon 12, has also been cloned. HDAC9Δexon15 is 1027 amino acids long and lacks a region within the catalytic domain adjacent to the active site. This isoform does not naturally conserve the ORF of HDAC9 and appears to have undergone RNA editing as shown below.

```
                          exon14           exon16
HDAC9         5'-aaa tgt gag -exon15- gt gat gac-3'
HDAC9Δexon15 5'-aaa tgt                     gac gac-3'
                  K   C                      D   D
```

Additionally, exon 4 may be deleted, which introduces an in-frame stop codon and results in a truncated protein of 95 amino acids. Exon 4 may also be alternatively spliced at either end as indicated in Table 2.

HDAC9 possesses histone deacetylase activity and represses basal transcription

In order to determine whether HDAC9 possesses histone deacetylase activity, an in-vitro assay was performed using anti-FLAG immunoprecipitated HDAC9. As shown in Figure 5A, HDAC9 contains about 25% activity of HDAC1 and 50% activity of HDACs 4 and 5. Higher activities of HDAC9 have been detected towards peptide substrates derived from histone H3 than from histone H4 (Figure 7A and B).

It has been previously established that HDACs repress transcription when tethered to DNA as Gal4 fusion proteins. This effect is also observed with HDAC9, HDAC9ΔCD and HDAC9CD alone (Figure 5B). A $GAL4^{uas}$x5-TK-Luciferase reporter gene was transiently transfected into 293T cells together with the expression vectors for the indicated GAL4 fusion proteins. The repressive effects of various GAL4DBD-HDAC9 fusions were partially relieved by the addition of the HDAC inhibitor Trichostatin A, including that of HDAC9ΔCD. Although HDAC9ΔCD lacks an HDAC domain it has been shown to bind other HDACs directly and via co-repressors, and this may be reflected in its ability to repress in a TSA sensitive manner. These results indicate that HDAC9 is able to use Histone H4 as a substrate and functions, at least in part, as a component of the repression machinery of the cell.

Expression patterns of HDAC9 in normal human tissue and cancer cells

Under normal circumstances, the expression of HDAC9 isoforms which contain the catalytic domain appears to be tightly regulated, with the protein only found at high levels in adult and foetal brain, and to a lesser extent skeletal muscle, testes, bone marrow, thymus, spleen, $CD14^{+ve}$, and $CD19^{+ve}$ cells (Figure 8). However, inspection of RT-PCR data reveals that HDAC9 is specifically expressed in TEL-AML1 positive and negative pre-B cell acute lymphoblastic leukaemia samples, and B cell lymphoma cell lines and patient samples (Figure 8A and B). HDAC9 is not expressed in various acute myeloid leukaemia cell lines, with the exception of acute monocytic leukaemia, possibly reflecting a common progenitor for B cell and monocytic lineages. HDAC9 is also found at lower levels in the erythroleukemia cell line, HEL. The leukaemic cell lines, which over-express HDAC9, also over-express the isoforms lacking exon7 and exon12 relative to normal levels (see Figures 8B and C). In one the CLL patient sample that has been stimulated with SAC/IL2, there is a clear change in the relative expression of full length and ΔExon12 isoforms (Figure 8C). There is evidence that the addition of HDAC inhibitors can alter the cellular distribution of HDACs presumably through interfering with protein-protein interactions as well blocking the catalytic site with respect to HDAC9, such findings should be considered in the context of various isoforms and their function. HDAC9 isoforms lacking functional domains encoded by exons 7 or 12 may prove to be an important factor in the pathogenesis of some hematalogical malignancies.

In-vitro and *in-vivo* interactions of HDAC9

Given the expression patterns of the HDAC9 gene in the normal and malignant cells deriving from the B cell lineage, we examined whether its products could interact with any of the proteins whose function has been implicated in B-cell malignancies (Fig. 9). As expected, HDAC9 was found to interact with Diffuse Large Cell Lymphoma and pre-B cell acute lymphoblastic leukaemia associated transcriptional repressors BCL-6 and TEL, respectively. HDAC9 also interacted with PLZF, HDAC 4 and to a lesser degree class I HDACs 1 and 3. In addition HDAC9 was found to interact with the nuclear receptor co-repressors mSin3A and B and N-CoR, whose activities have been implicated in the mechanism underlying pathogeneses of several human cancers.

Fig. 10 shows the cellular localization of various HDAC9 isoforms when visualized with anti-FLAG antibody and an antibody to the N-terminal epitope present in all of the HDAC9 isoforms. Note that HDAC9Δexon7 is completely excluded from the nucleus. To further analyse the interactions of HDAC9 isoforms in vivo, we performed the co-immunofluorescence assays (Fig. 11-13). When HDAC9 was co-expressed with BCL-6, PLZF and N-CoR, it was shown to colocalize with them *in-vivo* and to influence their wild type localization towards a pattern seen for a given HDAC9 isoform. These data indicated those in vitro interactions between the HDAC9 and a number of other proteins (see above and Fig. 9) can also be observed in vivo. For example, when HDAC9Δexon7 was co-expressed with a given interaction partner there was a dramatic change in its cellular localization, with HDAC9Δexon7 recruiting it to the cytoplasm (see Fig. 10).

The references mentioned herein are all expressly incorporated by reference.

Table 1
Human HDAC9 exon/intron splice junctions

Exon	Frame	Size	5' Splice Donor	Intron	Size	3' Splice Acceptor
1	-	109	TCTAAGCCAGgtttaattggtt	1	407	gtttttcctcagATGGGGTGGC
2	1	63	ATCAGCTCAGgtaagatcctct	2	88956	ctggttctttagTGGATGTGAA
3	2	242	GCATATCAAGgtagcaaatgct	3a	4812	tcttctcgcaagTTGCAACAGG
3	2			3b		aagttgcaacagGAACTTCTAG
4	1	142	GGACGAGAAAgtaagaggcacc	4c		
4	1		AGGCACCAGGgtaaacgatgga	4d	1014	tgtgtatttcagGGGCAGTGGC
5	2	127	TCTGGTACACgtatgttcagtg	5	2265	tgtcttttctagGGCTGCCCAC
6	3	122	CGAAAAACTGgtaagttggttt	6	35320	ctcaatccccagCCTCTGAGCC
[illegible]	2	132	GAGGTGACAGgtaattgaggac	7	5145	aatattttcagAATCCTCAGT
8	2	116	TCATGCCGAGgtaagaccctta	8	9928	ttttttaacagCAAATGGTTT
[illegible]	1	123	CCAGCTCAATgtaagtcattgc	9	2991	ttctcaacacagGCTTCGAATT
10	1	214	CTTGTAGCTGgtaattcattat	10	467	tttttttttcagGTGGAGTTCC
11	2	218	CATGAACAAAgtaagcctccaa	11	17529	actctcttctagCTGCTTTCGA
[illegible]	1	264	TATGCAACAGgtaataggcaaa	12	58741	tcttggcaacagCCTTTCCTGG
13	1	178	TCTGCAACTGgtaggaatccct	13	21243	cttgtcttaaagGAATTGCCTA
14	2	134	TAAATGTGAGgtaatccagaat	14	13018	attttcttgcagCGAATTCAAG
15	1	121	ATACTCCTAGgtctgtacgggc	15	4828	cttactgtatagGTGATGACTC
16	2	50	TGGACTTGGGgtaagtacaagt	16	26189	ctgtttgctcagGTGGACAGTG
[illegible]	1	108	AGAGCTGAAGgtgaggtccggg	17	35708	ttgttttcacagAATGGGTTTG
18	1	56	CCACAGCCATgtaagtaccagg	18	244	tctattccgcagGGGGTTCTGC
19	3	88	TGTAGATCTGgtatgtattcct	19	5921	atttccctgtagGATGTTCACC
[illegible]	1	120	CCCAAATGAGgttcggtttatt	20	313	ttctcttcccagGTTGGAACAG
21	1	98	AAGCATTCAGgttggtacttct	21	38480	tttactgtgcagGACCATCGTG
22	3	119	ACGGCAAAATgtaagtacctct	22	61212	gtattatggtagGTTTTGGTCA
23	2	134	AGGAAATGAGgtaaaaaagtaa	23	18203	ctattcttgcagCTGGAGCCAC
24	1	85	GAAATTCAAAgtatgtctttaa	24	21575	tgtttttcctagGCAAGTATTG
25	2	148	AAGACAGCAGgtatgaatccaa	25	20069	ttatttttacagAACTGCTGGT
26	3	40				

Consensus splice donor and acceptor sequences between exons (uppercase) and introns (lowercase) are underlined. Exons highlighted in ■ have the potential to be spliced out in-frame. Δ exon 7 and 12 cDNAs have already been detected. The suffixes a, b, c and d refer to the alternative splice sites at ether end of exon 4. i.e.

```
                         a        b
cacactctcatgtctttctcttctcgcaagTTGCAACagGAACTTCTAGCCATAAAACAGCAACAAGAACTCCTAGAAAA
GGAGCAGAAACTGGAGCAGCAGAGGCAAGAACAGGAAGTAGAGAGGCATCGCAGAGAACAGCAGCTTCCTCCTCTCAGAG
GCAAAGATAGAGGACGAGAAAgtAAGAGGCACCAGGgtaaacgatggactctctttcctcatcgttagctgatcattatt
                     c             d
```

Areas shown in *italics* at either end of Exon4 refer to alternatively splices regions of the exon that have been detected.

Table 2

Genomic organization of HDAC9 and its MITR isoform

Exon	Clone	Accession #	Location on BAC	Relative position
MITR/ HDAC9 5'				
1	CTA-317M2	AC002433	43921-44029	1-110
2	"	"	44434-44499	517-580
3	"	"	133456-133699	89536-89779
4	"	"	138520-138661	94600-94742
5	"	"	139691-139817	95771-95898
6	CTB-180O1	AC002124	2364-2485	98163-98283
7	"	"	37086-37937	133605-133737
8	"	"	43083-43198	138882-138998
9	"	"	53127-53249	148926-149049
10	"	"	56241-56454	152040-152254
11	"	"	56922-57139	152721-152939
12a	"	"	74669-77298	170468-173098
HDAC9 3'				
12b	CTB-180O1	AC002124	74669-74932	170468-170732
13	CTB-13P7	AC002088	35119-35296	231839-232017
14	"	"	56544-56677	253260-253394
15	"	"	69696-69816	266412-266533
16	"	"	74645-74694	271361-271411
17	CTA-264L19	AC002410	15451-15558	297600-297708
18	"	"	51267-51322	333416-333472
19	"	"	51567-51654	333716-333801
20	"	"	57573-57692	339722-339842
21	"	"	58006-58103	340155-340253
22	RP5-1194E15	AC004994	79771-79643	378733-378852
23	"	"	18434-18301	440064-440198
24	GS1-465N13	AC004744	85480-85390	458401-458486
25	"	"	63816-63669	480061-480209
26	"	"	43602-43430	500278-500449

Claims:

1. An isolated polypeptide having histone deacetylase activity, wherein the polypeptide has the amino acid sequence of the polypeptide in Figure 2.

2. An isolated polypeptide having histone deacetylase activity, wherein the polypeptide has at least 80% amino acid sequence identity with the polypeptide in Figure 2.

3. The isolated polypeptide of claim 2, wherein the polypeptide has at least 90% amino acid sequence identity with the polypeptide in Figure 2 and comprises all of part of the amino acid sequence as set out between amino acids 1009 to 1069 of the polypeptide of Figure 2.

4. An isolated polypeptide having histone deacetylase activity, wherein the polypeptide is encoded by a nucleic acid sequence capable of hybridising under stringent conditions to the nucleotide sequence of Figure 2, or a complement thereof, and wherein the nucleic acid sequence comprises a portion able to hybridise under stringent conditions to a nucleic acid sequence encoding amino acids 1009 to 1069 of the polypeptide of Figure 2, or a complement thereof.

5. An isolated polypeptide which is variant, fragment or active portion of a polypeptide according to any one of the preceding claims, wherein said variant, fragment or active portion has histone deacetylase activity.

6. The isolated polypeptide according to claim 5, wherein the polypeptide comprises all of part of the

amino acid sequence as set out between amino acids 1009 to 1069 of the polypeptide sequence of Figure 2.

7. The isolated polypeptide of claim 5 which is a variant in which one or more of the exons shown in Figure 2 is deleted.

8. The isolated polypeptide variant of claim 7 in which exon 7, exon 12 and/or exon 15 are deleted.

9. The isolated polypeptide of any one of the preceding claims which is joined to a coupling partner.

10. The isolated polypeptide of claim 9, wherein the coupling partner is an effector molecule, a label, a drug, a toxin or a carrier or transport molecule.

11. An isolated nucleic acid molecule encoding a polypeptide of any one of the preceding claims.

12. An isolated nucleic acid molecule having a sequence as shown in Figure 2 and which encodes a polypeptide having histone deacetylase activity.

13. An expression vector comprising a nucleic acid of claim 11 or claim 12, operably linked to control sequences to direct its expression.

14. A host cell transformed with the expression vector of claim 13.

15. A method of producing a histone deacetylase polypeptide of any one of claims 1 to 10, the method comprising culturing the host cells of claim 14 and

isolating the polypeptide thus produced.

16. A composition comprising a histone deactylase 9 polypeptide of any one of claims 1 to 10.

17. A composition comprising a nucleic acid molecule of claim 11 or claim 12.

18. Use of a histone deacetylase 9 (HDAC9) polypeptide of any one of claims 1 to 10, or a nucleic acid molecule encoding the polypeptide, for screening for candidate compounds which (a) share a histone deacetylase 9 biological activity or (b) bind to the histone deacetylase 9 polypeptide or (c) inhibit a biological activity of a histone deacetylase 9 polypeptide.

19. The use of claim 18, wherein the biological activity of the HDAC9 polypeptide is activity as a histone deacetylase.

20. A method of identifying a compound which is capable of modulating an activity of a histone deacetylase 9 polypeptide (HDAC9) of any one of claims 1 to 10, the method comprising:

(a) contacting at least one candidate compound with the HDAC9 polypeptide under conditions in which the candidate compound and HDAC9 polypeptide are capable of interacting;

(b) determining in an assay for a HDAC9 activity whether the candidate compound modulates the activity; and,

(c) selecting a candidate compound which modulates an activity of the HDAC9 polypeptide.

21. The method of claim 20, wherein the modulating is binding of the candidate compound to the HDAC9 polypeptide.

22. The method of claim 20, wherein the modulating is the candidate compound inhibiting an activity of the HDAC9 polypeptide.

23. The method of claim 22, wherein the activity is histone deacetylase activity.

24. The method of claim 23, which comprises contacting the HDAC9 polypeptide with a histone substrate in the presence or absence of the candidate compound and determining the amounts of acetylated and deacetylated histone substrate present after reaction with the HDAC9 polypeptide.

25. The method of claim 24, wherein the histone substrate is histone 3 or histone 4.

26. The method of claim 24 or claim 25, wherein the amounts of acetylated and deacetylated histone substrate is determined using antibodies which are capable of specifically binding to acetylated or deacetylated histone substrate.

27. The method of claim 24 or claim 25, wherein the histone substrate has a label which is releasable from the substrate by the HDAC9, so that the amounts of acetylated and deacetylated histone substrate are determined by measuring the amount of label released from the substrate by the action of the HDAC9 polypeptide.

28. The method of claim 27, wherein the label is a radioactive label or a fluorescent label.

29. A method of identifying a compound which is capable of inhibiting histone deacetylase 9 (HDAC9) polypeptide of any one of claims 1 to 10, the method comprising:

(a) contacting at least one candidate compound and the HDAC9 polypeptide in the presence of a substrate for HDAC9 under conditions in which the candidate compound, HDAC9 polypeptide and HDAC9 substrate are capable of interacting;

(b) determining whether the candidate compound inhibits the activity of the HDAC9 polypeptide in reacting with the substrate; and,

(c) selecting the candidate compound which inhibits the activity of the HDAC9 polypeptide on the substrate.

30. The method of claim 29, wherein the method employs a cell based reporter assay comprising (a) cells which are capable of producing HDAC9 polypeptide fused to a nucleic acid binding domain and (b) a reporter construct comprising (i) a promoter having a nucleic acid sequence that can be bound by the nucleic acid binding domain, thereby to activate the promoter, and (ii) a reporter under the transcriptional control of the promoter.

31. The method of claim 30, wherein the presence of active HDAC9 polypeptide causes deacetylation of core histones associated with the nucleic acid binding domain of the reporter and down regulation of promoter activity.

32. The method of claim 30 or claim 31, wherein the method is employed to identify candidate compounds that inhibit HDAC9 activity and counteract the down regulation

of the promoter.

33. The method of any one of claims 30 to 32, wherein the nucleic acid binding domain is a GAL4 domain, and the HDAC9-GAL4 protein which is capable of interacting with a GAL4 nucleic acid binding site of the promoter, so that HDAC9 polypeptide activity causes deacetylation of core histones associated with a DNA surrounding GAL4 binding site, thereby causing down regulation of promoter activity.

34. A method which comprises, having identified a candidate compound by the method of any one of claims 20 to 33, the step of manufacturing the compound and/or formulating it in a composition.

35. An antibody which is capable of specifically binding to a histone deacetylase 9 (HDAC9) polypeptide of any one of claims 1 to 10.

36. Use of an antibody of claim 35 in an assay to detect and quantify the presence of HDAC9 polypeptide, or an isoform thereof.

37. Use of an antibody of claim 35 for purifying s HDAC9 polypeptide.

38. Use of an antibody of claim 35 as an inhibitor of a HDAC9 polypeptide.

39. A method for the diagnosis or prognosis of cancer, the method comprising determining the presence or amount of histone deacetylase 9 (HDAC9) polypeptide of any one of claims 1 to 10, or an isoform thereof, or nucleic acid

encoding a HDAC9 polypeptide, in a sample from a patient.

40. The method of claim 39, wherein the cancer is a leukaemia.

41. The method of claim 40, wherein the leuakaemia is TEL-AML1 positive and negative pre-B cell acute lymphoblastic leukaemia or B cell lymphoma

42. The method of any one of claims 39 to 41, wherein the method comprises:

(a) determining the presence or amount of HDAC9 polypeptide in a sample from a patient, by measuring an activity of the HDAC9 polypeptide or its presence in a binding assay; or,

(b) determining the presence of HDAC9 nucleic acid using a probe capable of hybridising to the HDAC9 nucleic acid;

(c) using PCR involving one or more primers based on a HDAC9 nucleic acid sequence to determine whether the HDAC9 transcript is present in a sample from a patient.

43. The method of any one of claims 39 to 42, wherein the method comprises the steps of:

(a) contacting a sample obtained from the patient with a solid support having immobilised thereon binding agent having binding sites specific for HDAC9 polypeptide or HDAC9 nucleic acid;

(b) contacting the solid support with a labelled developing agent capable of binding to unoccupied binding sites, bound HDAC9 polypeptide or nucleic acid or occupied binding sites; and,

(c) detecting the label of the developing agent specifically binding in step (b) to obtain a value representative of the presence or amount of the HDAC9 polypeptide or nucleic acid in the sample.

44. Use of an inhibitor of histone deacetylase 9 (HDAC9) for the preparation of a medicament for the treatment of cancer.

45. The use of claim 44, wherein the cancer is a leukaemia.

46. The use of claim 45, wherein the leukaemia is TEL-AML1 positive and negative pre-B cell acute lymphoblastic leukaemia or B cell lymphoma.

1/17

```
AA287983 :  120 ggccttggagaagggtacaatataaatattgcctggacaggtggccttgatcctcccatgggagatgttgagtaccttgaagcattcaggaccatcgtgaagcctgtgg-caaagagttt  119
hHDAC9   : 2737 ggccttggagaagggtacaatataaatattgcctggacaggtggccttgatcctcccatgggagatgttgagtaccttgaagcattcaggaccatcgtgaagcctgtggccaaagagttt 2856

AA287983 :  120 gatccagacatggtcttagtatctgctggatttgatgcattggaaggccacacccctcctctaggagggtacaaa-gtgacggcaaaat 207
hHDAC9   : 2857 gatccagacatggtcttagtatctgctggatttgatgcattggaaggccacacccctcctctaggagggtacaaa-gtgacggcaaaatgttttggtcatttgacgaagcaattgatgac 2975
AA405905 :  370                                           gatgcattggaaggccacacccctcctctagggtttttcgaaagtgacggcaaaatgttttggtcatttgacgaagcaattgatgac  284

AA405905 :  283 attggctgatggacgtgtggtgttggctctagaaggaggacatgatctcacagccatctgtgatgcatcagaagcctgtgtaaatgcccttctaggaaatgagctggagccacttgcaga  164
hHDAC9   : 2976 attggctgatggacgtgtggtgttggctctagaaggaggacatgatctcacagccatctgtgatgcatcagaagcctgtgtaaatgcccttctaggaaatgagctggagccacttgcaga 3095

AA405905 :  163 agatattctcctccaaagcccgaatatgaatgctgttatttctttacagaagatcattgaaattcaaag 95
hHDAC9   : 3096 agatattctccaccaaagcccgaatatgaatgctgttatttctttacagaagatcattgaaattcaaagcaagtattggaagtcagtaaggatggtggctgtgccaaggggctgtgctct 3215
T06037   :    1                                                   tctttacagaagatcattgaaattcaaagcaagtattggaagtcagtaaggatggtggctgtnccaaggggctgtcctct   80

T06037   :   81 ggctggtgctcagttgcaagaggagacagagaccgtttctnccctggcctccctaacagtggatgtggaacagcccctttnctcaggaagacagcagaactgctggtgagcctatggaaga  200
hHDAC9   : 3216 ggctggtgctcagttgcaagaggagacagagaccgtttctgccctggcctccctaacagtggatgtggaacagcccctttgctcaggaagacagcagaactgctggtgagcctatggaaga 3335

T06037   :  201 ggagccagccttgtnaagtgccaagtcccccctctnatatttcctgtgtgtgacatcattgtgtatccccccaccccagnaccctcagacatgatctttgtctgctgactgggtggcacag  320
hHDAC9   : 3336 ggagccagccttgtgaagtgccaagtcccccctctgatatttcctgtgtgtgacatcattgtgtatccccccaccccagtaccctcagacatg-tc-ttgtctgctgcctgggtggcacag 3453
BE929866 :    1                                                    acatcattgggtatccccccaccccagtaccctcagacatg-tc-ttgtctgctgcctgggtggcacag   67

T06037   :  321 aattcaatngaacataaacaactgggcanaaaatt 355
hHDAC9   : 3454 -attcaatggaacataaac-actgggcacaaaattctgaacagcagcttcacttgttctttggatggacttgaaagggcattaaagattccttaaacgtaaccgctgtgattctagagtt 3571
BE929866 :   68 -attcaatggaacatgaac-actgggcacaaaattctgaacagcagcttcacttgttctttggatggacttgaaagggcattaaagattccttaaacgtaaccgctgtgattctagagtt  185

BE929866 :  186 acagtaaaccacgattggaggaaactgcttccagcatgcttttaatatgctgggtgacccactcctagacaccaagtttgaactagaaacattcagtacagcactagatattgttaattt  305
hHDAC9   : 3572 acagtaaaccacgattggaagaaactgcttccagcatgcttttaatatgctgggtgacccactcctagacaccaagtttgaactagaaacattcagtacagcactagatattgttaattt 3691

BE929866 :  306 cagaagctatgacagccagtgaaattttggcaaaacctgagacatagtcattcctgacattctgatcagctttttttggggtaatt  392
hHDAC9   : 3692 cagaagctatgacagccagtgaaattttggcaaaacctgagacatagtcattcctgacattctgatcagctttttttggggtaatt 3778
```

Fig. 1

```
                                                                  ggggaagagaggcacagacacagataggagaagggcaccggctggag
ccacttgcaggactgagggtttttgcaacaaaaccctagcagcctgaagaactctaagccag--- Intron 1 ---atggggtggctggacgagagcagctcttggctcagcaaaga
                                                                     407bp

      ATG CAC AGT ATG ATC AGC TCA G--- Intron 2 ---TG GAT GTG AAG TCA GAA GTT CCT GTG GGC CTG GAG CCC ATC TCA CCT TTA GAC CTA      26
HDAC9 M   H   S   M   I   S   S   V      88956 bp       D   V   K   S   E   V   P   V   G   L   E   P   I   S   P   L   D   L

AGG ACA GAC CTC AGG ATG ATG ATG CCC GTG GTG GAC CCT GTT GTC CGT GAG AAG CAA TTG CAG CAG GAA TTA CTT CTT ATC CAG CAG CAG      56
R   T   D   L   R   M   M   M   P   V   V   D   P   V   V   R   E   K   Q   L   Q   Q   E   L   L   L   I   Q   Q   Q

CAA CAA ATC CAG AAG CAG CTT CTG ATA GCA GAG TTT CAG AAA CAG CAT GAG AAC TTG ACA CGG CAG CAC CAG GCT CAG CTT CAG GAG CAT      86
Q   Q   I   Q   K   Q   L   L   I   A   E   F   Q   K   Q   H   E   N   L   T   R   Q   H   Q   A   Q   L   Q   E   H

ATC AAG--- Intron 3   ---TTG CAA CAG GAA CTT CTA GCC ATA AAA CAG CAA CAA GAA CTC CTA GAA AAG GAG CAG AAA CTG GAG CAG CAG      112
I   K      4813 bp       L   Q   Q   E   L   L   A   I   K   Q   Q   Q   E   L   L   E   K   E   Q   K   L   E   Q   Q

AGG CAA GAA CAG GAA GTA GAG AGG CAT CGC AGA GAA CAG CAG CTT CCT CCT CTC AGA GGC AAA GAT AGA GGA CGA GAA A--- Intron 4--      139
R   Q   E   Q   E   V   E   R   H   R   R   E   Q   Q   L   P   P   L   R   G   K   D   R   G   R   E   R    1029 bp

-GG GCA GTG GCA AGT ACA GAA GTA AAG CAG AAG CTT CAA GAG TTC CTA CTG AGT AAA TCA GCA ACG AAA GAC ACT CCA ACT AAT GGA AAA      168
    A   V   A   S   T   E   V   K   Q   K   L   Q   E   F   L   L   S   K   S   A   T   K   D   T   P   T   N   G   K

AAT CAT TCC GTG AGC CGC CAT CCC AAG CTC TGG TAC AC--- Intron 5 ---G GCT GCC CAC CAC ACA TCA TTG GAT CAA AGC TCT CCA CCC      194
N   H   S   V   S   R   H   P   K   L   W   Y   T       2265bp        A   A   H   H   T   S   L   D   Q   S   S   P   P

CTT AGT GGA ACA TCT CCA TCC TAC AAG TAC ACA TTA CCA GGA GCA CAA GAT GCA AAG GAT GAT TTC CCC CTT CGA AAA ACT G--- Intron      222
L   S   G   T   S   P   S   Y   K   Y   T   L   P   G   A   Q   D   A   K   D   D   F   P   L   R   K   T   A(E)  35320

6 ---CC TCT GAG CCC AAC TTG AAG GTG CGG TCC AGG TTA AAA CAG AAA GTG GCA GAG AGG AGA AGC AGC CCC TTA CTC AGG CGG AAG GAT      250
bp      S   E   P   N   L   K   V   R   S   R   L   K   Q   K   V   A   E   R   R   S   S   P   L   L   R   R   K   D

GGA AAT GTT GTC ACT TCA TTC AAG AAG CGA atg ttt gag gtg aca g--- Intron 7 ---AA TCC TCA GTC AGT AGC AGT TCT CCA GGC TCT      276
G   N   V   V   T   S   F   K   K   R   M   F   E   V   T   E      5145 bp       S   S   V   S   S   S   S   P   G   S

GGT CCC AGT TCA CCA AAC AAT GGG CCA ACT GGA AGT GTT ACT GAA AAT GAG ACT TCG GTT TTG CCC CCT ACC CCT CAT GCC GAG---Intro      304
G   P   S   S   P   N   N   G   P   T   G   S   V   T   E   N   E   T   S   V   L   P   P   T   P   H   A   E       9928

n 8   ---CAA ATG GTT TCA CAG CAA CGC ATT CTA ATT CAT GAA GAT TCC ATG AAC CTG CTA AGT CTT TAT ACC TCT CCT TCT TTG CCC AAC      332
bp       Q   M   V   S   Q   Q   R   I   L   I   H   E   D   S   M   N   L   L   S   L   Y   T   S   P   S   L   P   N

ATT ACC TTG GGG CTT CCC GCA GTG CCA TCC CAG CTC AAT--- Intron 9   ---GCT TCG AAT TCA CTC AAA GAA AAG CAG AAG TGT GAG ACG      358
I   T   L   G   L   P   A   V   P   S   Q   L   N       2991bp       A   S   N   S   L   K   E   K   Q   K   C   E   T

CAG ACG CTT AGG CAA GGT GTT CCT CTG CCT GGG CAG TAT GGA GGC AGC ATC CCG GCA TCT TCC AGC CAC CCT CAT GTT ACT TTA GAG GGA      388
Q   T   L   R   Q   G   V   P   L   P   G   Q   Y   G   G   S   I   P   A   S   S   S   H   P   H   V   T   L   E   G

AAG CCA CCC AAC AGC AGC CAC CAG GCT CTC CTG CAG CAT TTA TTA TTG AAA GAA CAA ATG CGA CAG CAA AAG CTT CTC GTA GCT G---Int      417
K   P   P   N   S   S   H   Q   A   L   L   Q   H   L   L   L   K   E   Q   M   R   Q   Q   K   L   L   V   A   G   467

ron 10---GT GGA GTT CCC TTA CAT CCT CAG TCT CCC TTG GCA ACA AAA GAG AGA ATT TCA CCT GGC ATT AGA GGT ACC CAC AAA TTG CCC      444
 bp         G   V   P   L   H   P   Q   S   P   L   A   T   K   E   R   I   S   P   G   I   R   G   T   H   K   L   P

CGT CAC AGA CCC CTG AAC CGA ACC CAG TCT GCA CCT TTG CCT CAG AGC ACG TTG GCT CAG CTG GTC ATT CAA CAG CAA CAC CAG CAA TTC      474
R   H   R   P   L   N   R   T   Q   S   A   P   L   P   Q   S   T   L   A   Q   L   V   I   Q   Q   Q   H   Q   Q   F

TTG GAG AAG CAG AAG CAA TAC CAG CAG CAG ATC CAC ATG AAC AAA--- Intron 11 ---CTG CTT TCG AAA TCT ATT GAA CAA CTG AAG CAA      500
L   E   K   Q   K   Q   Y   Q   Q   Q   I   H   M   N   K       17529 bp       L   L   S   K   S   I   E   Q   L   K   Q

CCA GGC AGT CAC CTT GAG GAA GCA GAG GAA GAG CTT CAG GGG GAC CAG GCG ATG CAG GAA GAC AGA GCG CCC TCT AGT GGC AAC AGC ACT      530
P   G   S   H   L   E   E   A   E   E   E   L   Q   G   D   Q   A   M   Q   E   D   R   A   P   S   S   G   N   S   T

AGG AGC GAC AGC AGT GCT TGT GTG GAT GAC ACA CTG GGA CAA GTT GGG GCT GTG AAG GTC AAG GAG GAA CCA GTG GAC AGT GAT GAA GAT      560
R   S   D   S   S   A   C   V   D   D   T   L   G   Q   V   G   A   V   K   V   K   E   E   P   V   D   S   D   E   D

GCT CAG ATC CAG GAA ATG GAA TCT GGG GAG CAG GCT GCT TTT ATG CAA CAG--- Intron 12 ---CCT TTC CTG GAA CCC ACG CAC ACA CGT      586
A   Q   I   Q   E   M   E   S   G   E   Q   A   A   F   M   Q   Q       61107 bp       P   F   L   E   P   T   H   T   R

HDAC9                                                     GTA ATA GGC AAA GAT TTA GCT CCA GGA TTT GTA ATT AAA      590
ΔCD                                                       V   I   G   K   D   L   A   P   G   F   V   I   K
Ex12
```

Fig. 2
(Part 1 of 2)

```
      GCG CTC TCT GTG CGC CAA GCT CCG CTG GCT GCG GTT GGC ATG GAT GGA TTA GAG AAA CAC CGT CTC GTC TCC AGG ACT CAC TCT TCC CCT   616
      A   L   S   V   R   Q   A   P   L   A   A   V   G   M   D   G   L   E   K   H   R   L   V   S   R   T   H   S   S   P
HDAC9
ΔCD   GTC ATT ATC TGA                                                                                                          593
Ex12  V   I   I   *
      GCT GCC TCT GTT TTA CCT CAC CCA GCA ATG GAC CGC CCC CTC CAG CCT GGC TCT GCA ACT G--- Intron 13---GA ATT GCC TAT GAC CCC   642
      A   A   S   V   L   P   H   P   A   M   D   R   P   L   Q   P   G   S   A   T   G      21243 bp      I   A   Y   D   P
      TTG ATG CTG AAA CAC CAG TGC GTT TGT GGC AAT TCC ACC ACC CAC CCT GAG CAT GCT GGA CGA ATA CAG AGT ATC TGG TCA CGA CTG CAA   672
      L   M   L   K   H   Q   C   V   C   G   N   S   T   T   H   P   E   H   A   G   R   I   Q   S   I   W   S   R   L   Q
      GAA ACT GGG CTG CTA AAT AAA TGT GAG--- Intron 14 ---CGA ATT CAA GGT CGA AAA GCC AGC CTG GAG GAA ATA CAG CTT GTT CAT TCT   698
      E   T   G   L   L   N   K   C   E      13018 bp      R   I   Q   G   R   K   A   S   L   E   E   I   Q   L   V   H   S
      GAA CAT CAC TCA CTG TTG TAT GGC ACC AAC CCC CTG GAC GGA CAG AAG CTG GAC CCC AGG ATA CTC CTA G--- Intron 15---GT GAT GAC   724
      E   H   H   S   L   L   Y   G   T   N   P   L   D   G   Q   K   L   D   P   R   I   L   L   G      4828 bp      D   D
      TCT CAA AAG TTT TTT TCC TCA TTA CCT TGT GGT GGA CTT GGG --- Intron 16---GTG GAC AGT GAC ACC ATT TGG AAT GAG CTA CGC TCG   750
      S   Q   K   F   F   S   S   L   P   C   G   G   L   G       26189 bp    V   D   S   D   T   I   W   N   E   L   R   S
      TCC GGT GCT GCA CGC ATG GCT GTT GGC TGT GTC ATC GAG CTG GCT TCC AAA GTG GCC TCA GGA GAG CTG AAG --- Intron 17---AAT GGG   776
      S   G   A   A   R   M   A   V   G   C   V   I   E   L   A   S   K   V   A   S   G   E   L   K       35708 bp      N   G
      TTT GCT GTT GTG AGG CCC CCT GGC CAT CAC GCT GAA GAA TCC ACA GCC AT---Intron 18 ---G GGG TTC TGC TTT TTT AAT TCA GTT GCA   802
      F   A   V   V   R   P   P   G   H   H   A   E   E   S   T   A   M     244 bp        G   F   C   F   F   N   S   V   A
      ATT ACC GCC AAA TAC TTG AGA GAC CAA CTA AAT ATA AGC AAG ATA TTG ATT GTA GAT CTG--- Intron 19 ---GAT GTT CAC CAT GGA AAC   828
      I   T   A   K   Y   L   R   D   Q   L   N   I   S   K   I   L   I   V   D   L      5918 bp        D   V   H   H   G   N
      GGT ACC CAG CAG GCC TTT TAT GCT GAC CCC AGC ATC CTG TAC ATT TCA CTC CAT CGC TAT GAT GAA GGG AAC TTT TTC CCT GGC AGT GGA   858
      G   T   Q   Q   A   F   Y   A   D   P   S   I   L   Y   I   S   L   H   R   Y   D   E   G   N   F   F   P   G   S   G
      GCC CCA AAT GAG--- Intron 20 ---GTT GGA ACA GGC CTT GGA GAA GGG TAC AAT ATA AAT ATT GCC TGG ACA GGT GGC CTT GAT CCT CCC   884
      A   P   N   E      313 bp        V   G   T   G   L   G   E   G   Y   N   I   N   I   A   W   T   G   G   L   D   P   P
      ATG GGA GAT GTT GAG TAC CTT GAA GCA TTC AG---Intron 21 ---G ACC ATC GTG AAG CCT GTG GCC AAA GAG TTT GAT CCA GAC ATG GTC   910
      M   G   D   V   E   Y   L   E   A   F   R      38480 bp      T   I   V   K   P   V   A   K   E   F   D   P   D   M   V
      TTA GTA TCT GCT GGA TTT GAT GCA TTG GAA GGC CAC ACC CCT CCT CTA GGA GGG TAC AAA GTG ACG GCA AAA T--- Intron 22---GT TTT   936
      L   V   S   A   G   F   D   A   L   E   G   H   T   P   P   L   G   G   Y   K   V   T   A   K   C      61212bp      F
      GGT CAT TTG ACG AAG CAA TTG ATG ACA TTG GCT GAT GGA CGT GTG GTG TTG GCT CTA GAA GGA GGA CAT GAT CTC ACA GCC ATC TGT GAT   956
      G   H   L   T   K   Q   L   M   T   L   A   D   G   R   V   V   L   A   L   E   G   G   H   D   L   T   A   I   C   D
      GCA TCA GAA GCC TGT GTA AAT GCC CTT CTA GGA AAT GAG--- Intron 23 ---CTG GAG CCA CTT GCA GAA GAT ATT CTC CAC CAA AGC CCG   992
      A   S   E   A   C   V   N   A   L   L   G   N   E      18203 bp      L   E   P   L   A   E   D   I   L   H   Q   S   P
      AAT ATG AAT GCT GTT ATT TCT TTA CAG AAG ATC ATT GAA ATT CAA A--- Intron 24---GC AAG TAT TGG AAG TCA GTA AGG ATG GTG GCT  1018
      N   M   N   A   V   I   S   L   Q   K   I   I   E   I   Q   S      21575 bp      K   Y   W   K   S   V   R   M   V   A
      GTG CCA AGG GGC TGT GCT CTG GCT GGT GCT CAG TTG CAA GAG GAG ACA GAG ACC GTT TCT GCC CTG GCC TCC CTA ACA GTG GAT GTG GAA  1048
      V   P   R   G   C   A   L   A   G   A   Q   L   Q   E   E   T   E   T   V   S   A   L   A   S   L   T   V   D   V   E
      CAG CCC TTT GCT CAG GAA GAC AGC AG---Intron 25 ---A ACT GCT GGT GAG CCT ATG GAA GAG GAG CCA GCC TTG TGA agtgccaagtccccc  1069
      Q   P   F   A   Q   E   D   S   R      20069 bp      T   A   G   E   P   M   E   E   E   P   A   L   *
      tctgatatttcctgtgtgtgacatcattgtgtatccccccaccccagtaccctcagacatgtcttgtctgctgcctgggtggcacagattcaatggaacataaacactgggcacaaaat
      tctgaacagcagcttcacttgttctttggatggacttgaaagggcattaaagattccttaaacgtaaccgctgtgattctagagttacagtaaaccacgattggaagaaactgcttcca
      gcatgcttttaatatgctgggtgacccactcctagacaccaagtttgaactagaaacattcagtacagcactagatattgttaatttcagaagctatgacagccagtgaaattttgggc
      aaaacctgagacatagtcattcctgacattctgatcagcttttttttggggtaatt
```

Fig. 2
(Part 2 of 2)

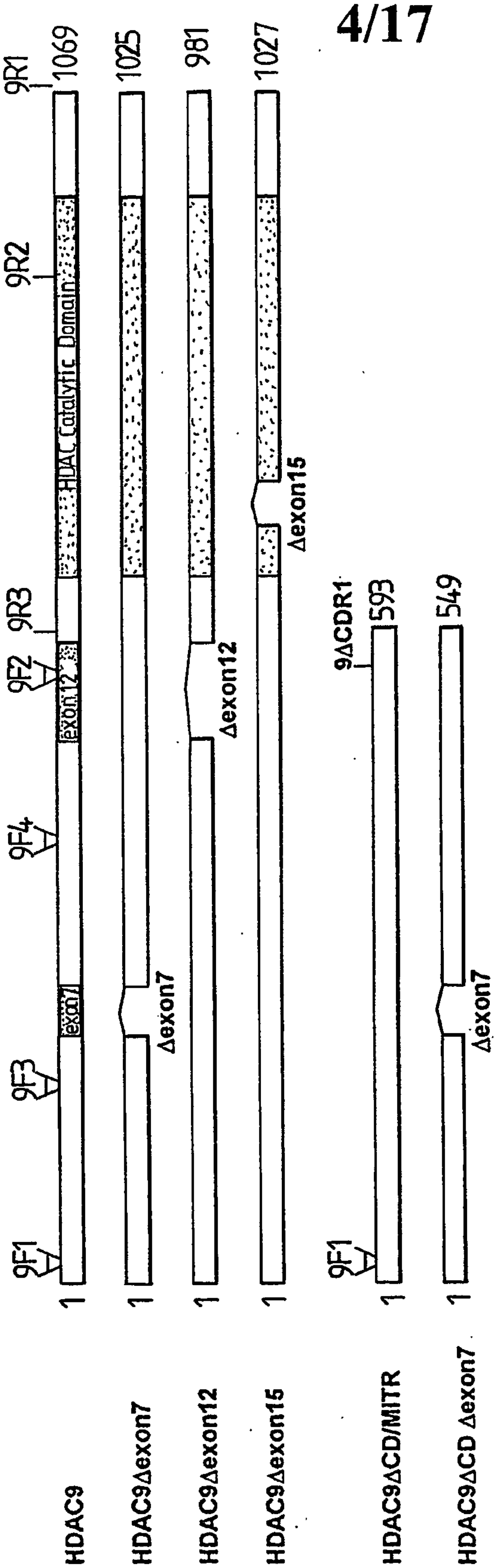

HDAC9
HDAC9Δexon7
HDAC9Δexon12
HDAC9Δexon15
HDAC9ΔCD/MITR
HDAC9ΔCD Δexon7
9F1
9F3
9F4
9F2
9R3
9R2
9R1
exon7
exon12
HDAC Catalytic Domain
1069
1025
981
1027
593
549
1
Δexon7
Δexon12
Δexon15
9ΔCDR1

Fig. 3

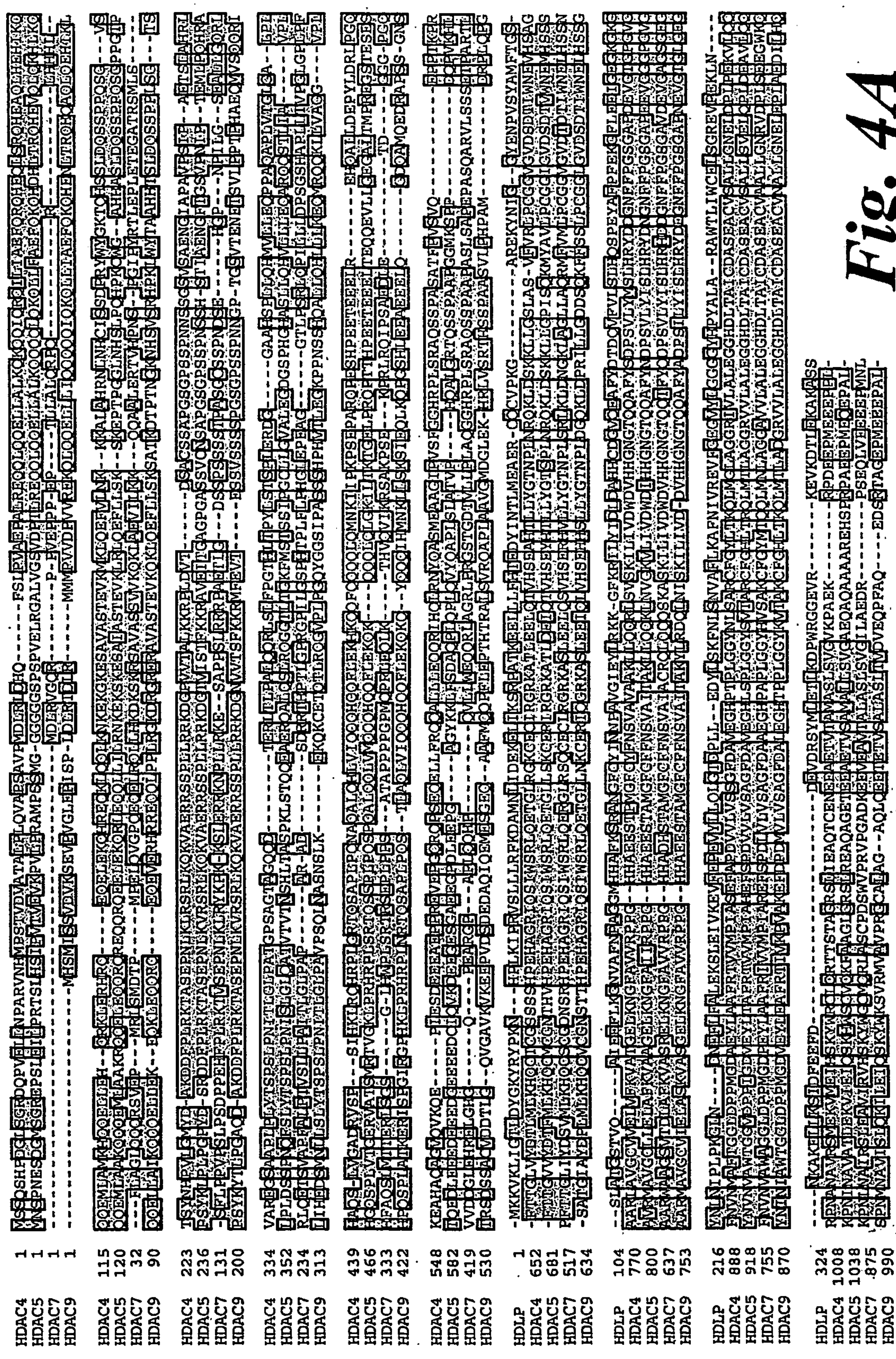


Fig. 4A

	Entire Protein		Catalytic Domain	
	Identity (%)	Similarity (%)	Identity (%)	Similarity (%)
HDAC9 vs HDAC1	26	45	26	45
HDAC9 vs HDAC4	56	70	70	83
HDAC9 vs HDAC5	57	71	72	84
HDAC9 vs HDAC7	43	54	66	77
HDAC9 vs HDAC6	17	28	N-term 39	56
			C-term 44	60

Fig. 4B

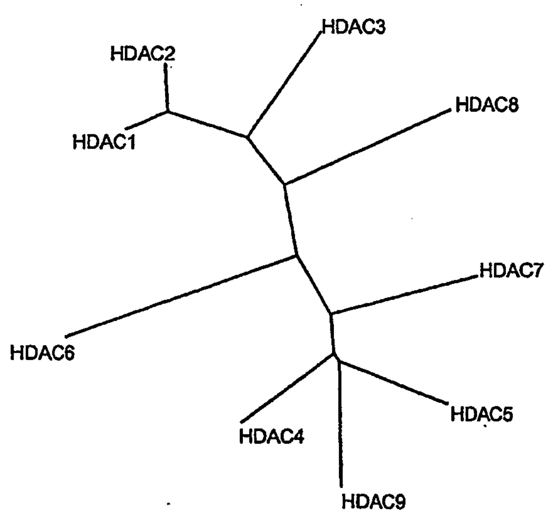


Fig. 4C

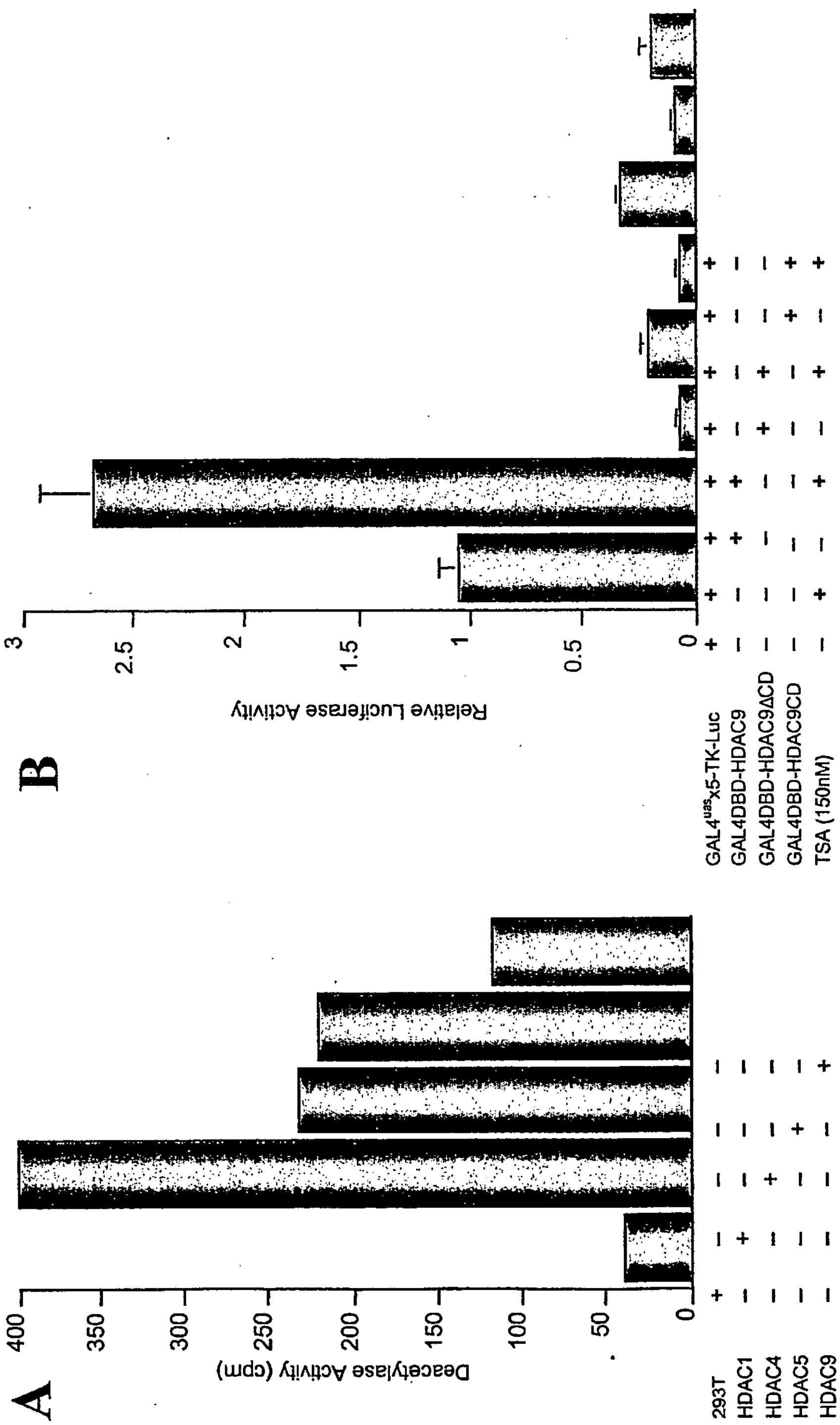

A
Deacetylase Activity (cpm)
0
50
100
150
200
250
300
350
400
293T
HDAC1
HDAC4
HDAC5
HDAC9
B
Relative Luciferase Activity
0
0.5
1
1.5
2
2.5
3
GAL4^{UAS}x5-TK-Luc
GAL4DBD-HDAC9
GAL4DBD-HDAC9ΔCD
GAL4DBD-HDAC9CD
TSA (150nM)

Fig. 5

8/17

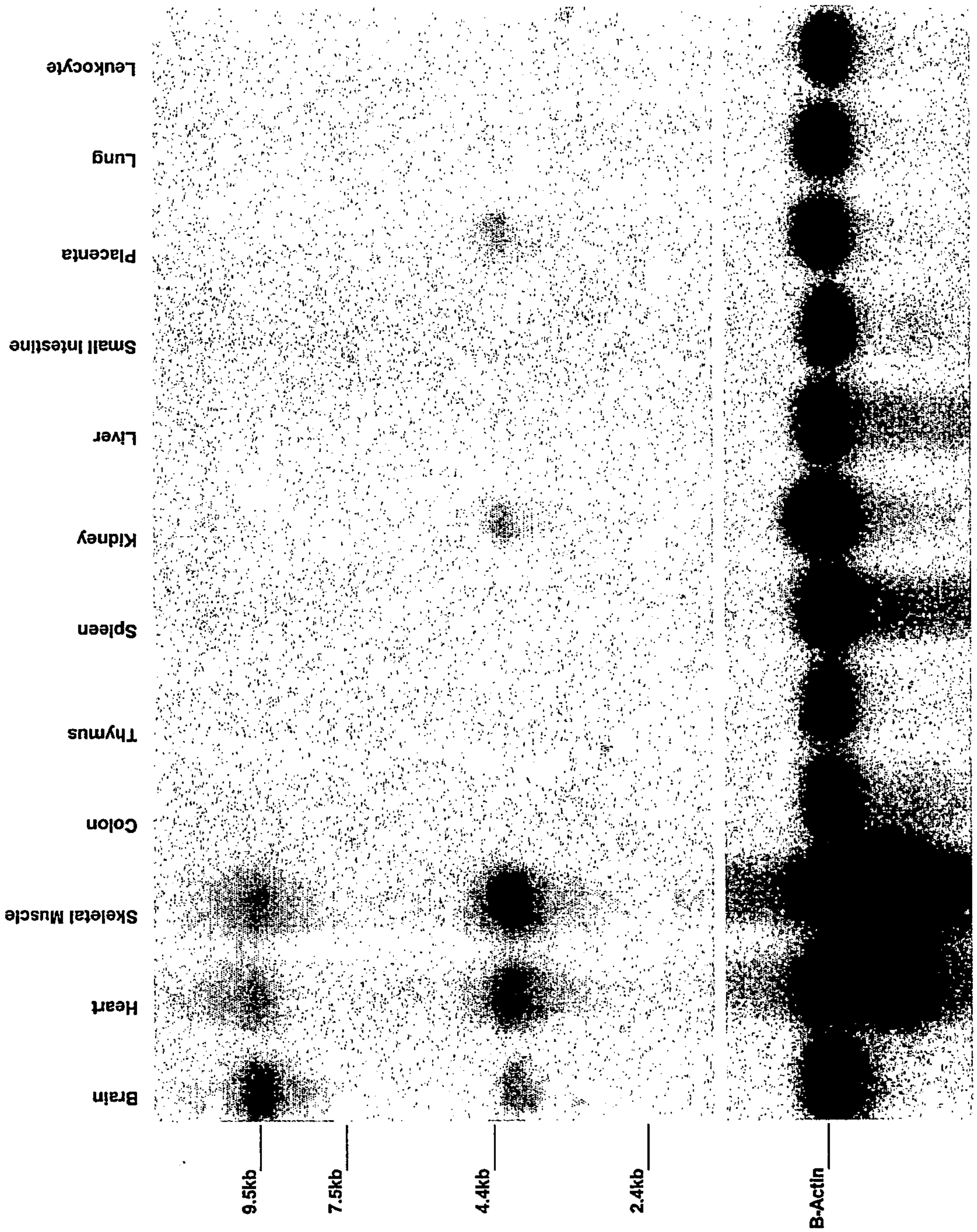

Leukocyte
Lung
Placenta
Small Intestine
Liver
Kidney
Spleen
Thymus
Colon
Skeletal Muscle
Heart
Brain
9.5kb
7.5kb
4.4kb
2.4kb
B-Actin

Fig. 6

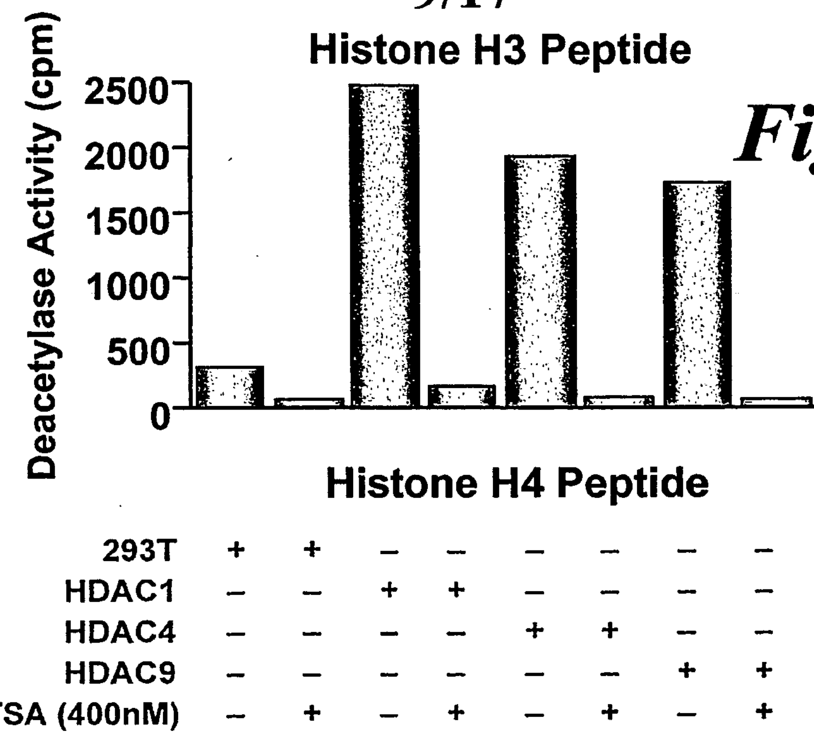
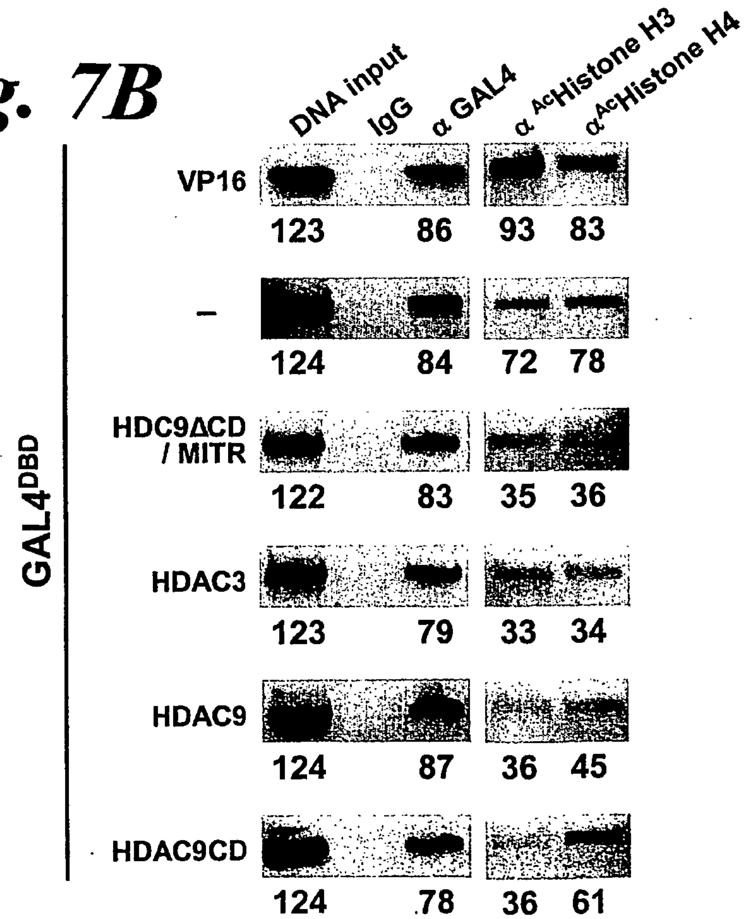

Histone H3 Peptide
Deacetylase Activity (cpm)
2500
2000
1500
1000
500
0
Histone H4 Peptide
293T + + – – – – – –
HDAC1 – – + + – – – –
HDAC4 – – – – + + – –
HDAC9 – – – – – – + +
TSA (400nM) – + – + – + – +

Fig. 7A

Fig. 7B


DNA input
IgG
α GAL4
α AcHistone H3
α AcHistone H4
GAL4DBD
VP16
123 86 93 83
–
124 84 72 78
HDC9ΔCD / MITR
122 83 35 36
HDAC3
123 79 33 34
HDAC9
124 87 36 45
HDAC9CD
124 78 36 61

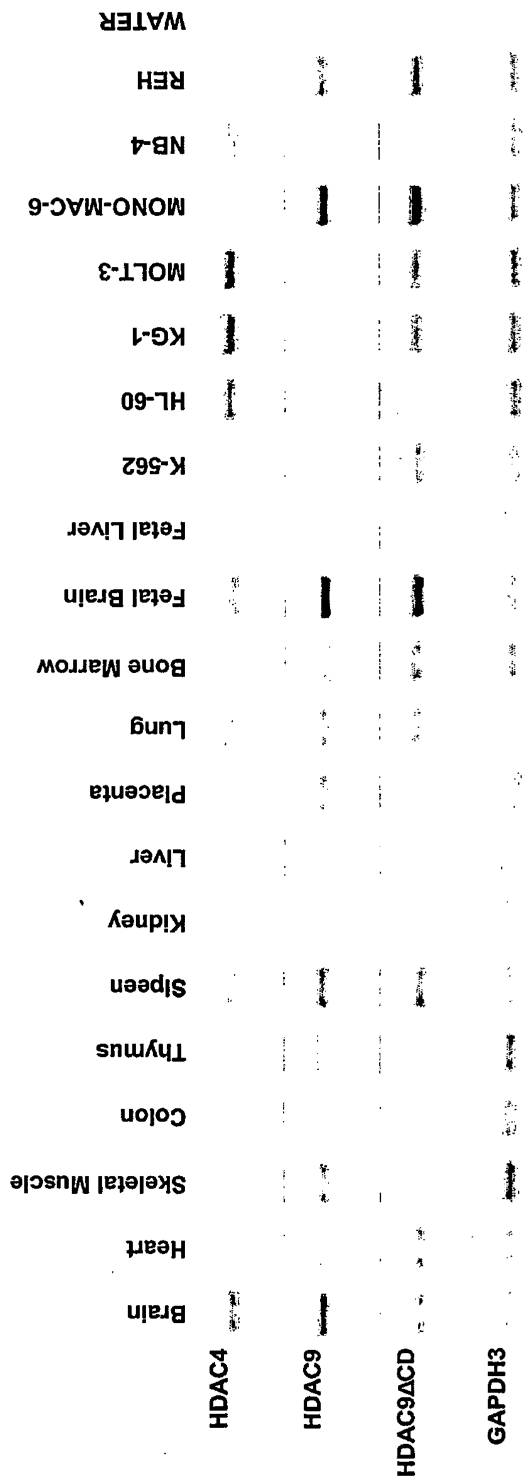

WATER
REH
NB-4
MONO-MAC-6
MOLT-3
KG-1
HL-60
K-562
Fetal Liver
Fetal Brain
Bone Marrow
Lung
Placenta
Liver
Kidney
Slpeen
Thymus
Colon
Skeletal Muscle
Heart
Brain
HDAC4
HDAC9
HDAC9ΔCD
GAPDH3

Fig. 8A

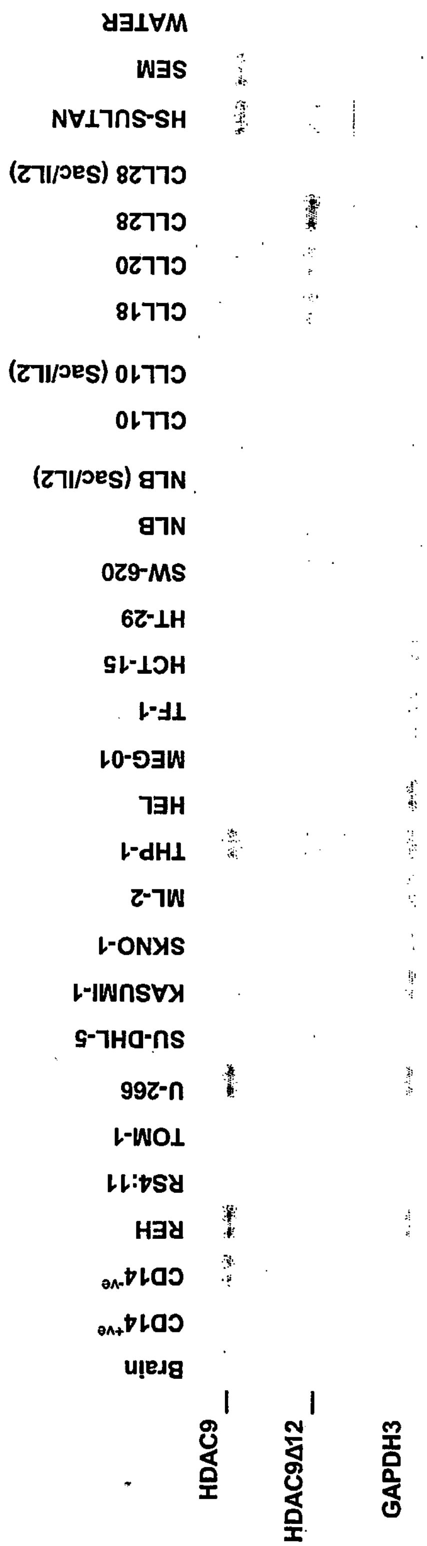

Brain
CD14+ve
CD14-ve
REH
RS4:11
TOM-1
U-266
SU-DHL-5
KASUMI-1
SKNO-1
ML-2
THP-1
HEL
MEG-01
TF-1
HCT-15
HT-29
SW-620
NLB
NLB (Sac/IL2)
CLL10
CLL10 (Sac/IL2)
CLL18
CLL20
CLL28
CLL28 (Sac/IL2)
HS-SULTAN
SEM
WATER
HDAC9
HDAC9Δ12
GAPDH3

Fig. 8B

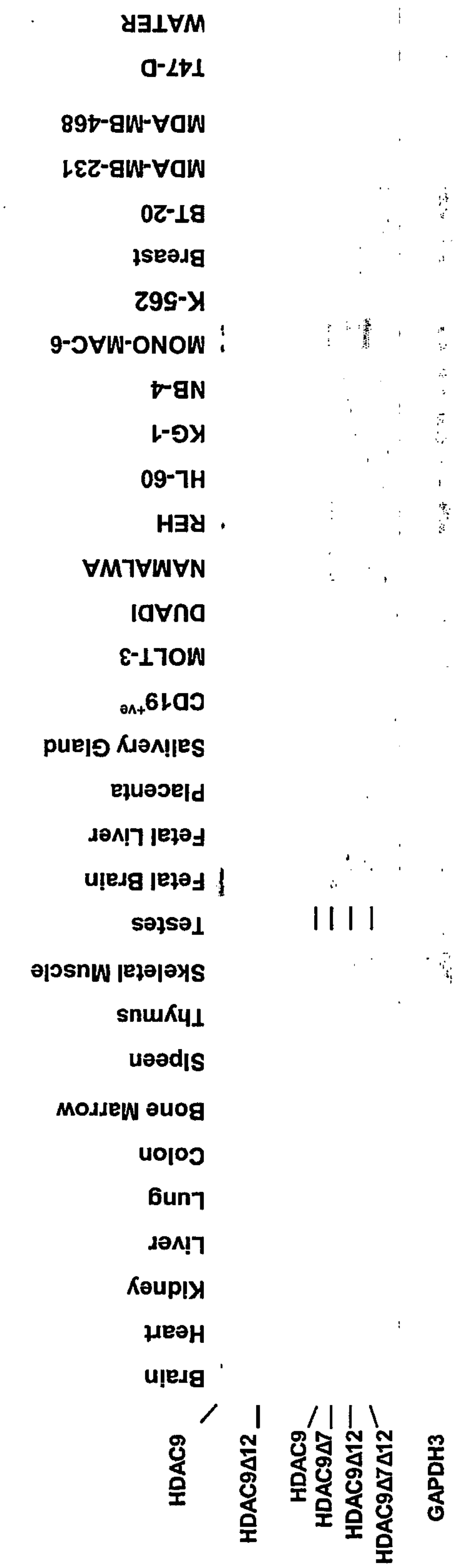

WATER
T47-D
MDA-MB-468
MDA-MB-231
BT-20
Breast
K-562
MONO-MAC-6
NB-4
KG-1
HL-60
REH
NAMALWA
DUADI
MOLT-3
CD19+ve
Salivery Gland
Placenta
Fetal Liver
Fetal Brain
Testes
Skeletal Muscle
Thymus
Slpeen
Bone Marrow
Colon
Lung
Liver
Kidney
Heart
Brain
HDAC9
HDAC9Δ12
HDAC9
HDAC9Δ7
HDAC9Δ12
HDAC9Δ7Δ12
GAPDH

Fig. 8C

13/17

Fig. 9

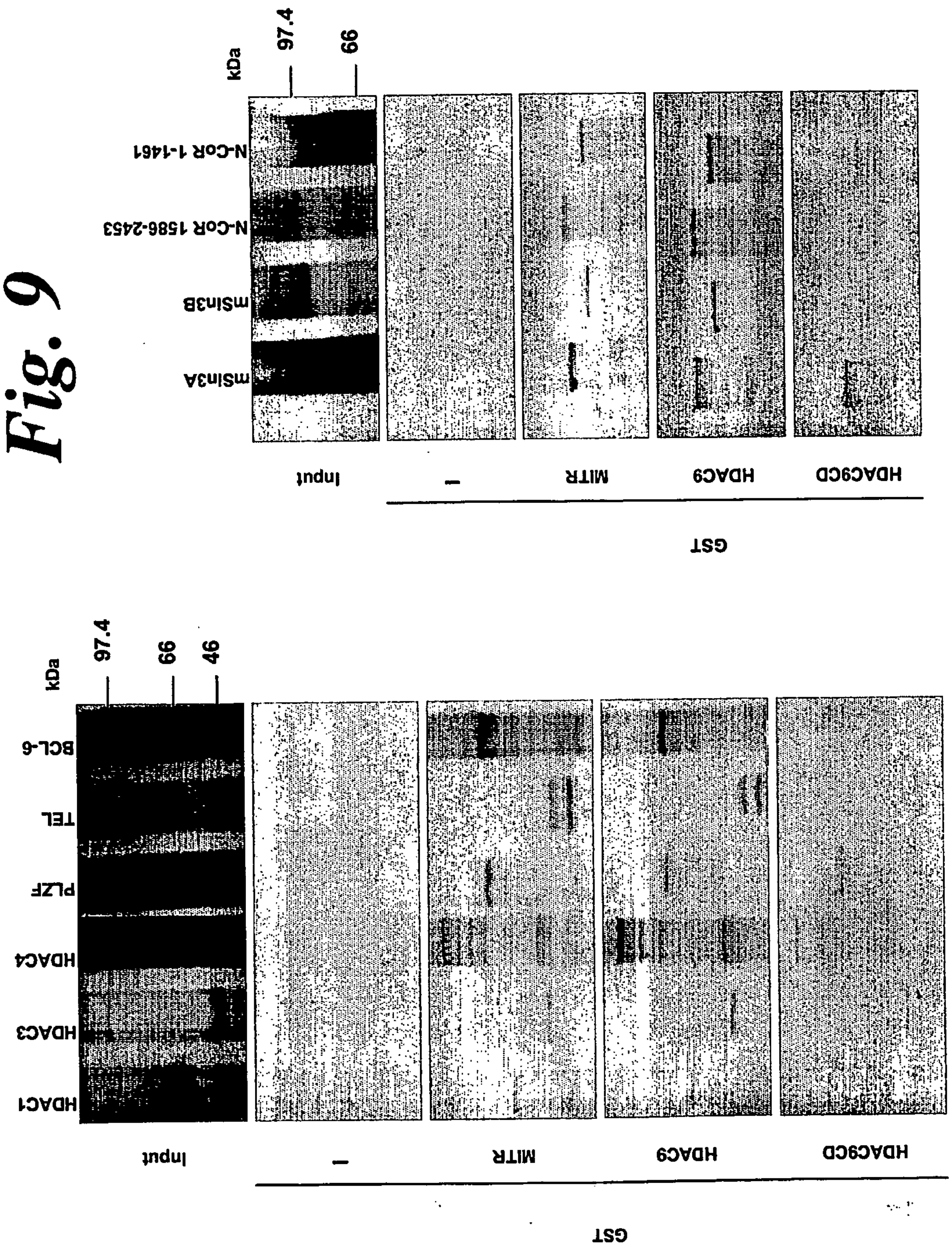
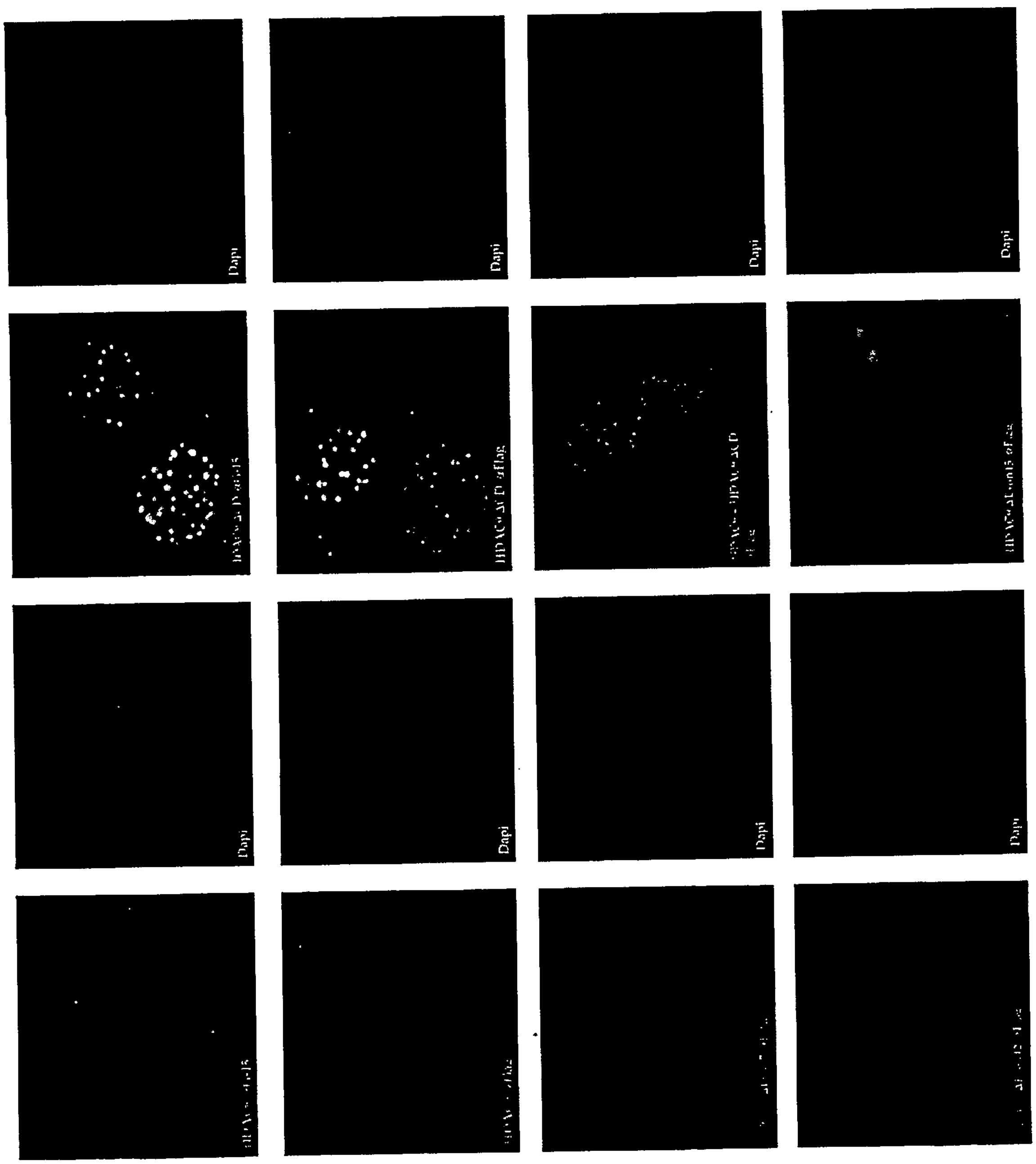

kDa
97.4
66
N-CoR 1-1461
N-CoR 1586-2453
mSin3B
mSin3A
Input
–
MITR
HDAC9
HDAC9CD
GST
kDa
97.4
66
46
BCL-6
TEL
PLZF
HDAC4
HDAC3
HDAC1
Input
–
MITR
HDAC9
HDAC9CD
GST

Fig. 10



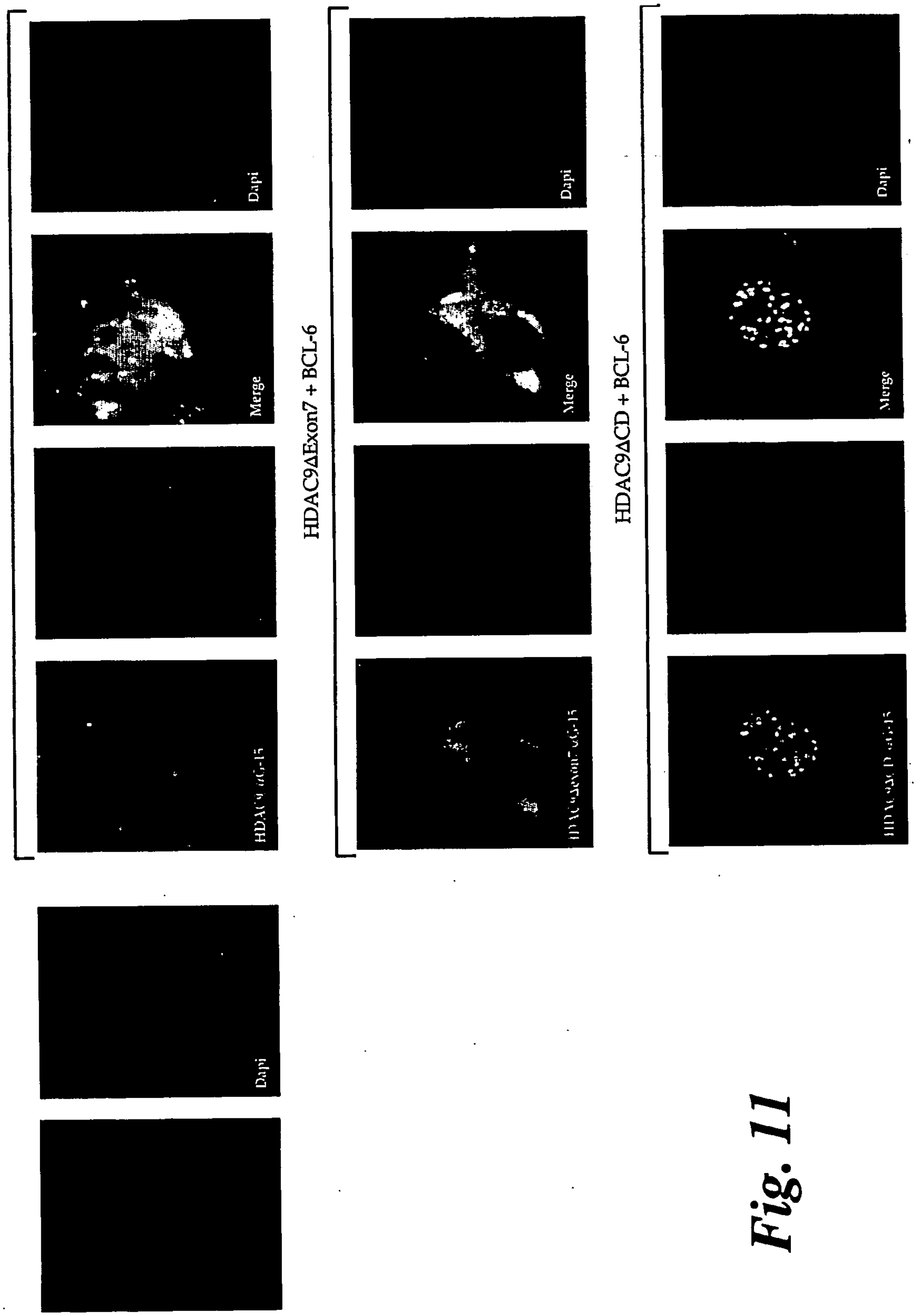

HDAC9 + BCL-6
HDAC9ΔExon7 + BCL-6
HDAC9ΔCD + BCL-6
Dapi
Merge
Dapi
Merge
Dapi
Merge
Dapi

Fig. 11

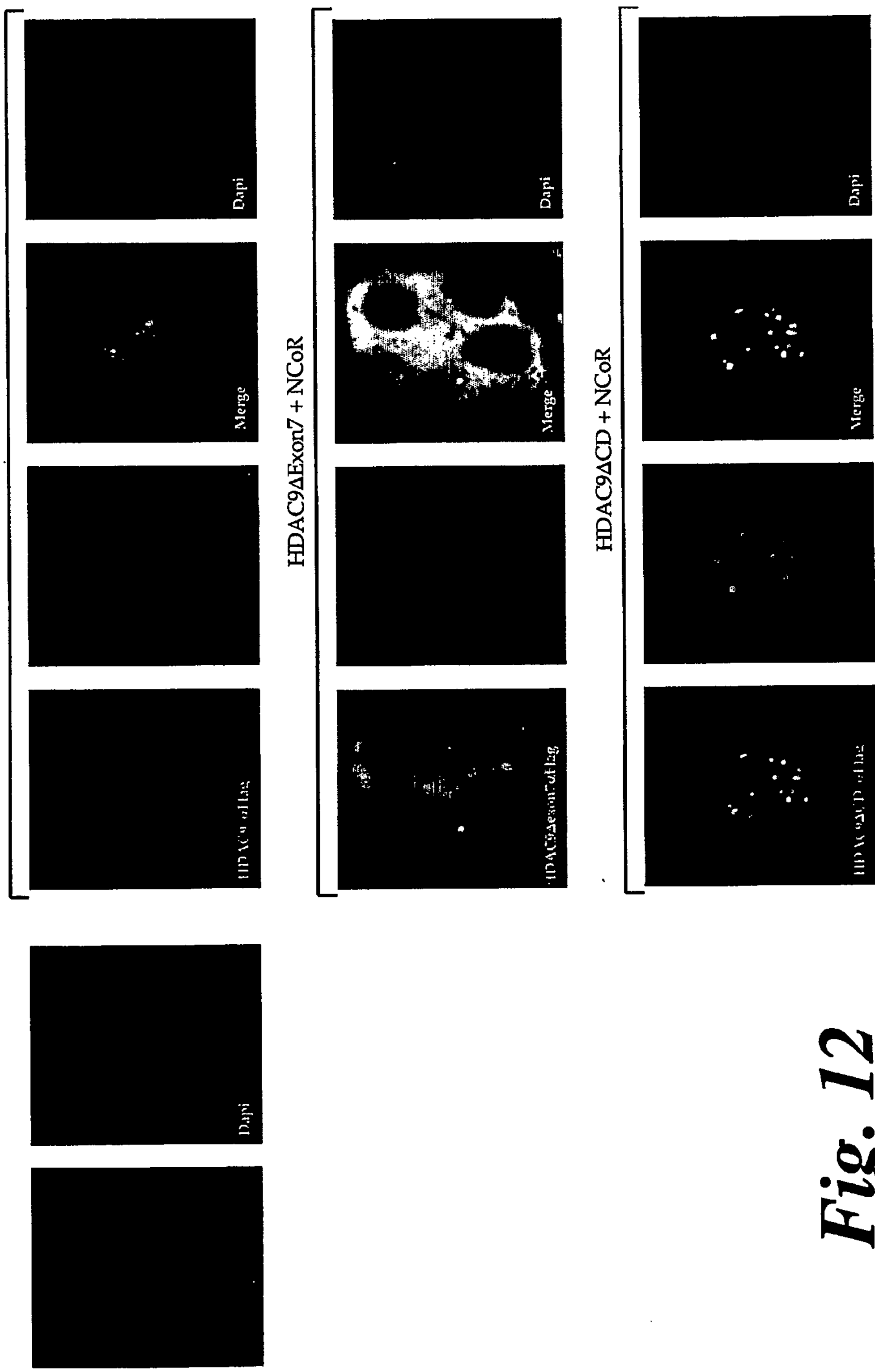

HDAC9 + NCoR
HDAC9ΔExon7 + NCoR
HDAC9ΔCD + NCoR
Dapi
Merge
Dapi
Merge
Dapi
Merge
Dapi

Fig. 12

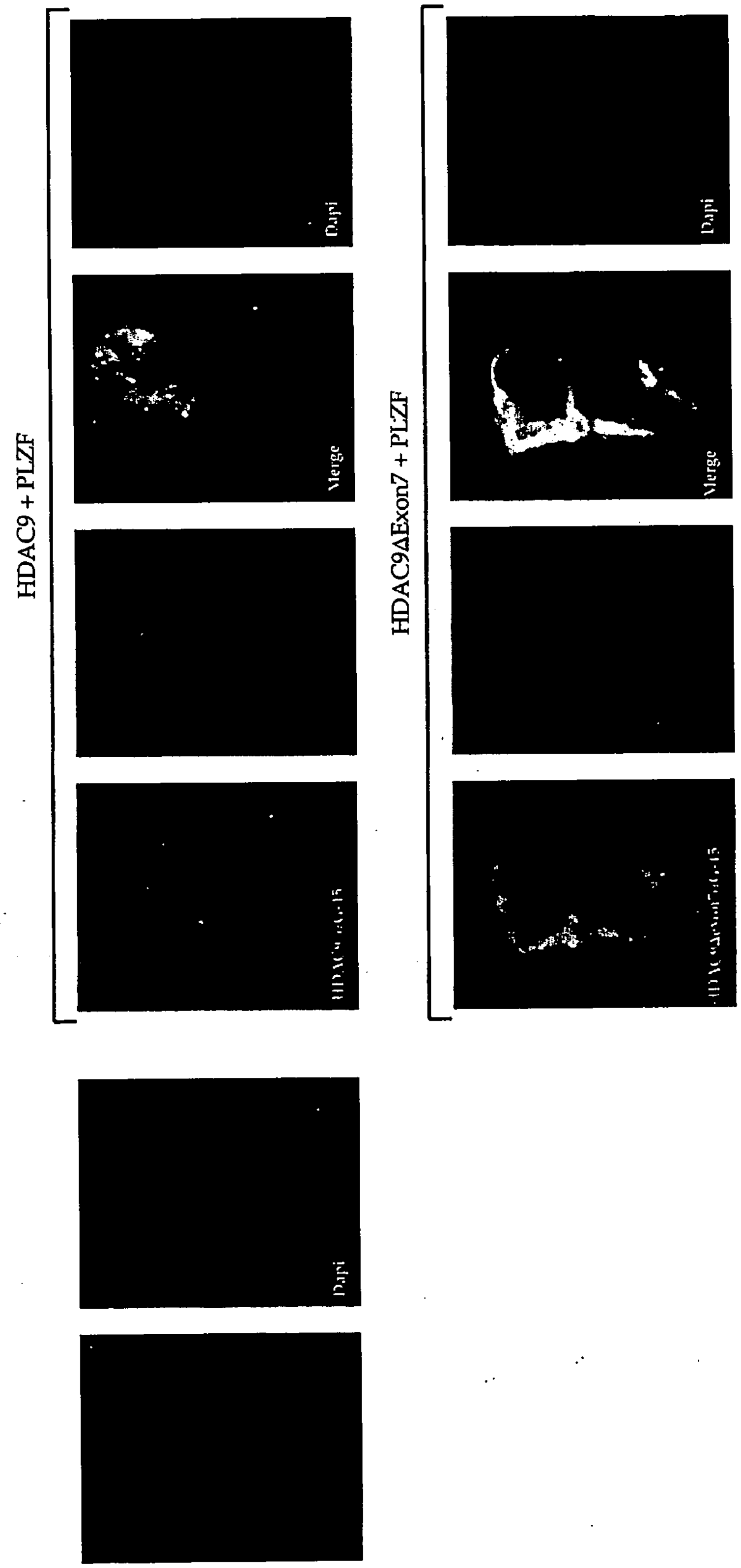


Fig. 13

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 April 2003 (10.04.2003)

PCT

(10) International Publication Number
WO 03/029451 A3

(51) International Patent Classification[7]: C12N 9/16, 15/55, 5/10, C12Q 1/34, 1/68, C07K 16/40, G01N 33/573, A61K 38/46, 48/00, A61P 35/00, 35/00

(21) International Application Number: PCT/GB02/04455

(22) International Filing Date: 2 October 2002 (02.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0123664.5 2 October 2001 (02.10.2001) GB

(71) Applicant *(for all designated States except US)*: **THE INSTITUTE OF CANCER RESEARCH** [GB/GB]; 123 Old Brompton Road, London, Greater London SW7 3RP (GB).

(71) Applicants and
(72) Inventors: **ZELENT, Arthur** [US/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB). **PETRIE, Kevin** [GB/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB). **GUIDEZ, Fabien** [FR/GB]; c/o The Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London, Greater London SW3 6JB (GB).

(74) Agents: **KIDDLE, Simon, J.** et al.; Mewburn Ellis, York House, 23 Kingsway, London, Greater London WC2B 6HP (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report: 21 August 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/029451 A3

(54) Title: HISTONE DEACETYLASE 9

(57) Abstract: The identification and cloning of histone deacetylase 9 (HDAC9) is disclosed, and in particular full length HDAC9 polypeptides and HDAC9 polypeptides which have deacetylase activity, and to nucleic acid molecules encoding these polypeptides. The uses of these polypeptides and nucleic acid molecules are disclosed, for example for screening for compounds that are capable of modulating a HDAC9 biological activity.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/04455

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N9/16 C12N15/55 C12N15/62 C12N5/10 C12Q1/34 C12Q1/68 C07K16/40 G01N33/573 A61K38/46 A61K48/00 A61P35/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ZHOU XIANBO ET AL: "Cloning and characterization of a histone deacetylase, HDAC9." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 98, no. 19, 11 September 2001 (2001-09-11), pages 10572-10577, XP002235661 September 11, 2001 ISSN: 0027-8424 cited in the application the whole document	5,7-10, 13-29, 34-38, 44-46
Y	figure 4	30-33
Y		39-43
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 March 2003	09/04/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Aslund, J

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WANG AUDREY H ET AL: "HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor." MOLECULAR AND CELLULAR BIOLOGY, vol. 19, no. 11, November 1999 (1999-11), pages 7816-7827, XP002235662 ISSN: 0270-7306 figure 1	5,6
Y	figure 5	30-33
Y	WO 01 42437 A (AXYS PHARM INC ;BUGGY JOSEPH (US)) 14 June 2001 (2001-06-14) the whole document	39-43
A	KAO H-Y ET AL: "ISOLATION OF A NOVEL HISTONE DEACETYLASE REVEALS THAT CLASS I AND CLASS II DEACETYLASES PROMOTE SMRT-MEDIATED REPRESSION" GENES AND DEVELOPMENT, COLD SPRING HARBOR LABORATORY PRESS, NEW YORK, US, vol. 14, 1 January 2000 (2000-01-01), pages 55-66, XP000981426 ISSN: 0890-9369 figure 3	18-34
A	YOSHIDA M ET AL: "POTENT AND SPECIFIC INHIBITION OF MAMMALIAN HISTONE DEACETYLASE BOTH IN VIVO AND IN VITRO BY TRICHOSTATIN A" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 265, no. 28, 5 October 1990 (1990-10-05), pages 17174-17179, XP000616087 ISSN: 0021-9258 the whole document	1-46
A	NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. XI. THE COMPLETE SEQUENCES OF 100 NEW CDNA CLONES FROM BRAIN WHICH CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 5, 1998, pages 277-286, XP000828191 ISSN: 1340-2838 See clone KIAA0744	1-46

-/--

INTERNATIONAL SEARCH REPORT

Internε Application No
PCT/GB 02/04455

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BERTOS N R ET AL: "CLASS II HISTONE DEACETYLASES: STRUCTURE, FUNCTION AND REGULATION" BIOCHEMISTRY AND CELL BIOLOGY. BIOCHIMIE ET BIOLOGIE CELLULAIRE, XX, XX, vol. 79, no. 3, 7 December 2000 (2000-12-07), pages 243-252, XP009004311 ISSN: 0829-8211	
E	WO 02 102984 A (RIFKIND RICHARD A ;SLOAN KETTERING INST CANCER (US); ZHOU XIANBO () 27 December 2002 (2002-12-27) the whole document	1-46
P,X	WO 02 36783 A (BAYER AG ;XIAO YONGHONG (US)) 10 May 2002 (2002-05-10) Nucleic acid seq Id no 6, Protein seq id no 7.	1-46
P,X	PETRIE KEVIN R ET AL: "Role of a novel Class II histone deacetylase in normal and leukemia-associated transcriptional repression." BLOOD, vol. 98, no. 11 Part 1, 16 November 2001 (2001-11-16), page 568a XP008015437 43rd Annual Meeting of the American Society of Hematology, Part 1;Orlando, Florida, USA; December 07-11, 2001, November 16, 2001 ISSN: 0006-4971 abstract	39-43

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

The claims have been renumbered as claims 1-46 in consequitive order for the purpose of this search report.

2. [X] Claims Nos.: 21-28, 32-34 (all partially)
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 21-28, 32-34 (all partially)

Present claims 21-28, 32-34 relate to a compound defined by reference to a desirable characteristic or property, namely the ability to bind to or modulate the activity of HDAC9. The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to an antibody directed to HDAC9.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Intern Application No
PCT/GB 02/04455

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 0142437	A	14-06-2001	AU	1959001 A	18-06-2001
			EP	1246905 A2	09-10-2002
			WO	0142437 A2	14-06-2001
WO 02102984	A	27-12-2002	WO	02102984 A2	27-12-2002
WO 0236783	A	10-05-2002	AU	1235002 A	15-05-2002
			WO	0236783 A2	10-05-2002